Satellites, Sensors, and Synthesis

Catharine van Ingen
Partner Architect
eScience Group, Microsoft Research
NICTA 25 October 2010

Science is Scaling

13-15 orders of Magnitude

From Dennis Baldocchi 2010

Eons, 10^9 yr
- Evolution of Life and the Formation of our Atmosphere

Geological Periods, 10^6 yr
- Evolution, Speciation, Extinction, Climate Regimes

Millennia, 10^3 yr
- Species migration

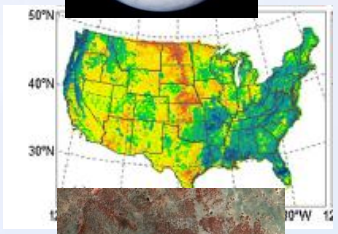
Century, 10^2 yr
- Succession, Mortality, Soil Formation

Decadal, 10^1 yr
- Competition, Gap-Replacement, Stand Dynamics
- Changes in Soil Organic Matter

Seasonal & Annual, 10^0 yr
- Net and Gross Primary Productivity
- Autotrophic and Heterotrophic Respiration and Decomposition
- Plant Acclimation
- Mineralization and Immobilization

Seconds/Hours/Day, 10^{-6} to 10^{-3} yr
- Photosynthesis, Respiration, Transpiration, Stomatal Conductance

Globe: 10,000 km (10^7 m)

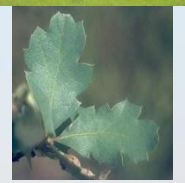


Continent: 1000 km (10^6 m)

Landscape: 1-100 km

Canopy: 100-1000 m

Plant: 1-10 m



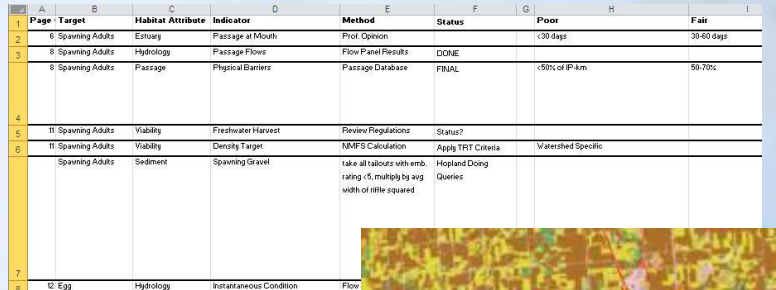
Leaf: 0.01-0.1 m

Stomata: 10^{-5} m

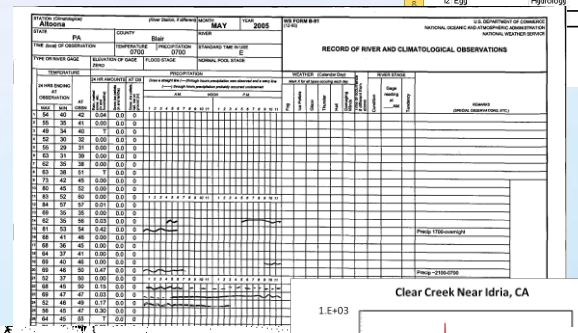
Bacteria/Chloroplast: 10^{-6} m

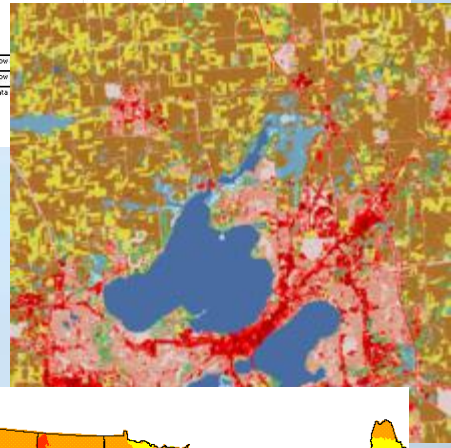
The Data is Out There

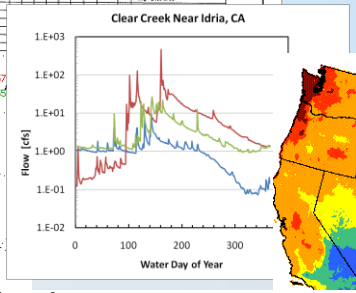
- National and International Datasets
  - USGS National Water Information System
  - NOAA National Climatic Data Center
  - FLUXNET Network
  - Satellite data (e.g. MODIS)
- Local Datasets
  - Local Agencies
  - Companies (e.g. Timber)
  - Ecology Organizations
  - Individual Researchers

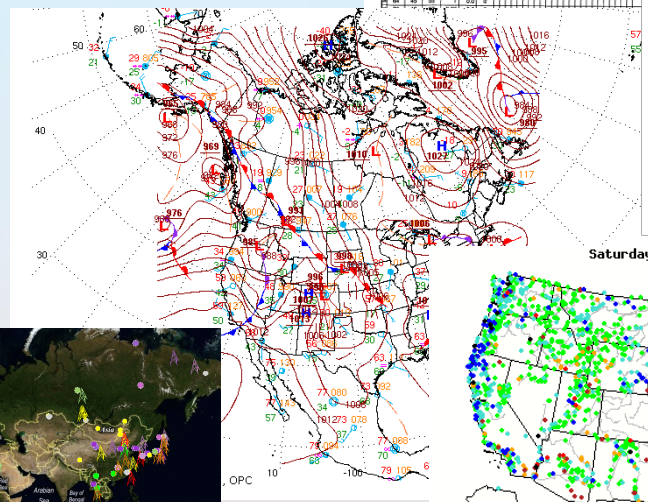


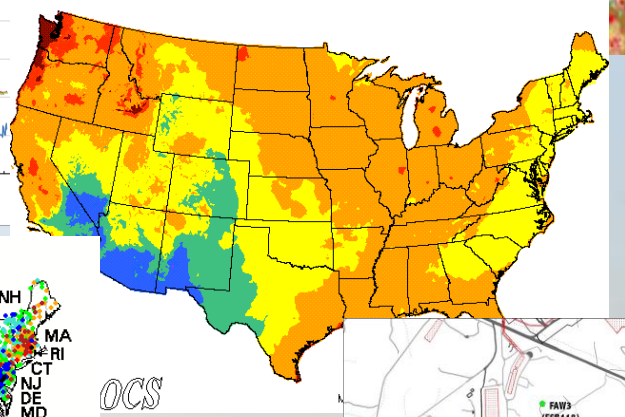
Page	Target	Habitat Attribute	Indicator	Method	Status	Poor	Fair
6	Spawning Adults	Estuary	Passage at Mouth	Prof. Opinion		<30 days	30-60 days
8	Spawning Adults	Hydrology	Passage Flows	Flow Panel Results	DONE		
8	Spawning Adults	Passage	Physical Barriers	Passage Database	FINAL	<50% of IP-km	50-70%
11	Spawning Adults	Viability	Freshwater Harvest	Review Regulations	Status?		
11	Spawning Adults	Viability	Density Target	NMFS Calculation	Apply TRT Criteria	Watershed Specific	
	Spawning Adults	Sediment	Spawning Gravel	take all tailouts with emb. rating <5, multiply by avg width of riffle squared	Hopland Doing Queries		
12	Egg	Hydrology	Instantaneous Condition	Flow			
			Redd Scour	Flow			
			Gravel Quality (Bulk)	Data			

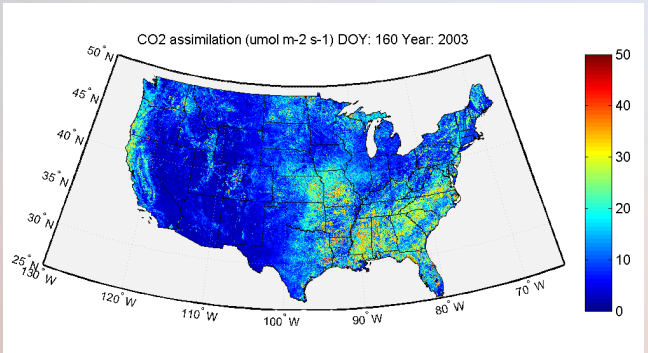


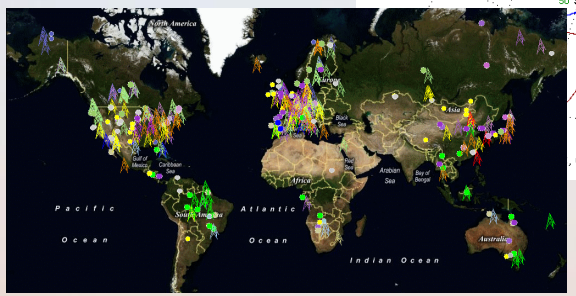


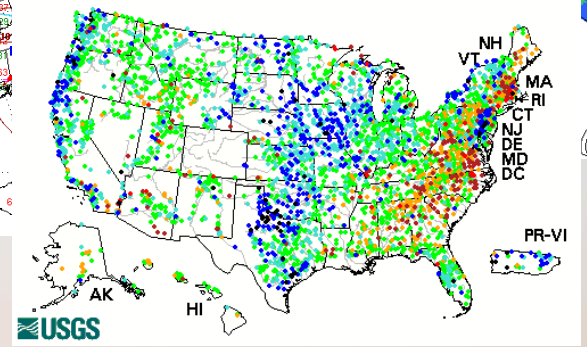


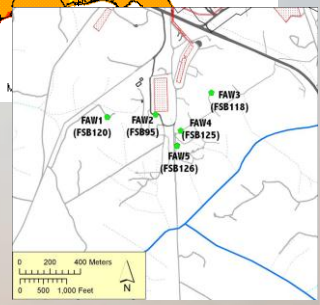










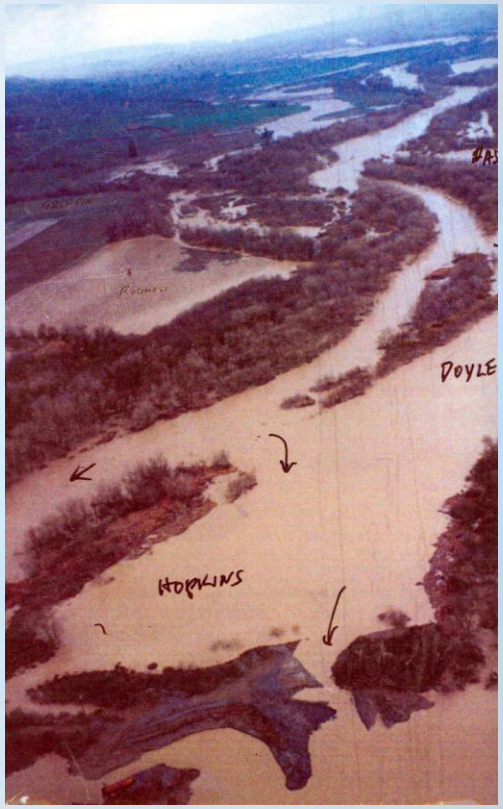


Data Variety – The Spice of Life

Manual Measurement

Automated Measurement

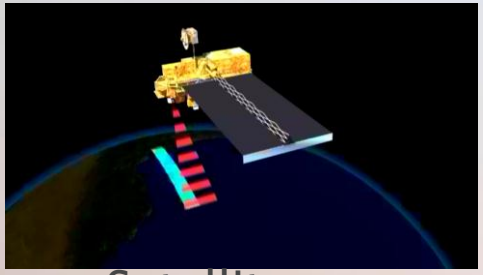
Sample Collection



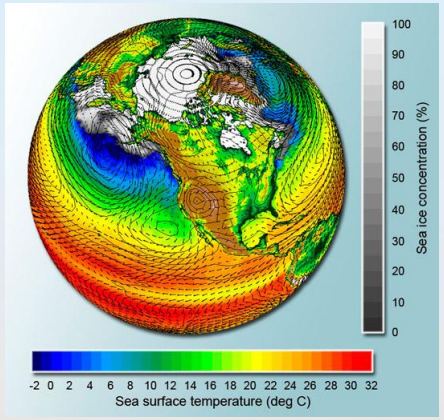
Historical Photographs

Typing

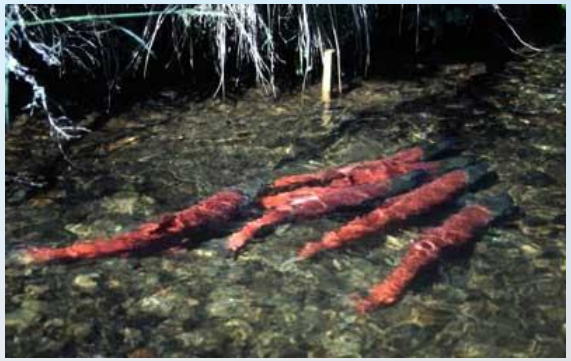
Aircraft Surveys



Satellite



Model Output



Counting

Relatively Ubiquitous Motes

Synthesis: BESS*

Big Science ! Hallelujah!
Big Science ! Yodelie Hoo!
Laurie Anderson

**Breathing Earth Science Simulator*

What is ET ?

- Evapotranspiration (ET) is the release of water to the atmosphere by evaporation from open water bodies and transpiration, or evaporation through plant membranes, by plants.
- Climate change isn't just about a change in temperature, it's also about a change in the water balance and hence water supply critical to human activity.

From Dr. Youngryel Ryu's science research proposal:

Evapotranspiration (*E*) is a major component of the terrestrial hydrological cycle (ca. 60% of precipitation) [*Trenberth, et al.*, 2007]. It controls land-atmosphere feedbacks and constitutes an important source of water vapor to the atmosphere [*Raupach*, 1998]. In turn, atmospheric water vapor is the most significant greenhouse gas and thus plays a fundamental role in weather and climate [*Held and Soden*, 2000]. Understanding *E* is important for socio-economic reasons, such as regulating available water for human use [*Brauman, et al.*, 2007]. Thus, there have been diverse efforts to regularly monitor *E* in a regional scale using satellite remote sensing imagery [*Anderson, et al.*, 2008; *Diak, et al.*, 2004; *Nishida, et al.*, 2003].

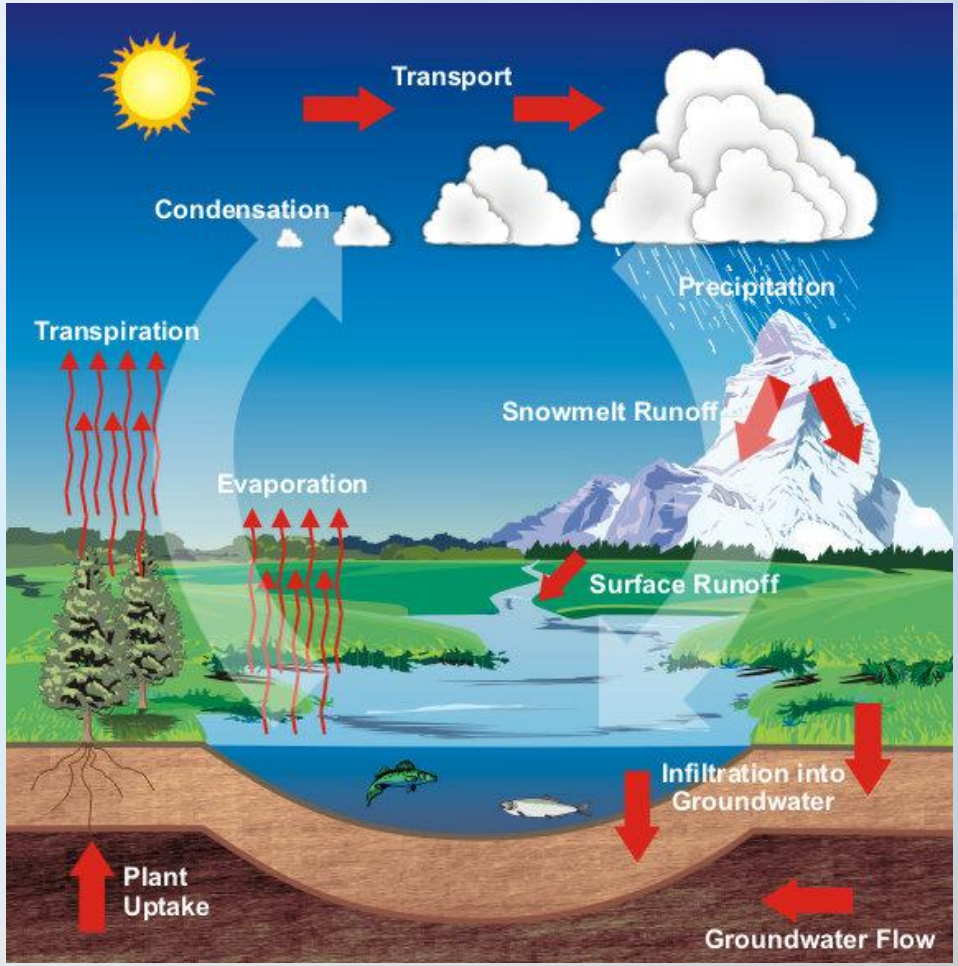


Image courtesy of the National Oceanic and Atmospheric Administration

Computing ET From Historical Sensor Data

$$ET = P - R - \frac{dS}{dt}$$

Simple Water Balance

ET: Evapotranspiration or release of water to the atmosphere by evaporation from open water bodies and transpiration by plants

P: Precipitation including snowfall

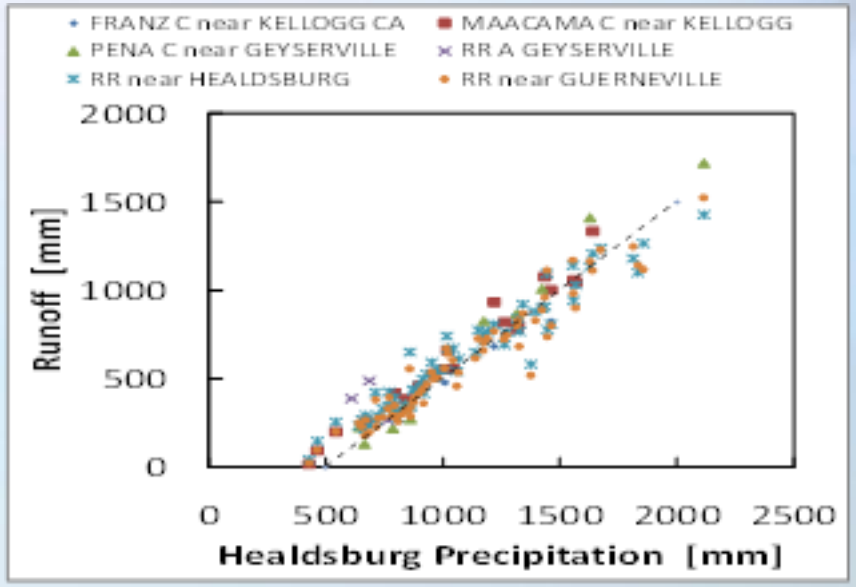
R: Surface runoff in streams and rivers

dS/dt: change in water storage over time such as increase in lakes or groundwater levels

P: http://www.ncdc.noaa.gov/oa/ncdc.html

R: http://waterdata.usgs.gov/nwis

- Easy to do (with a digital watershed)
- Long term trends only



In Mediterranean climates such as California, a long term equilibrium may exist. The ecosystem determines ET by soils and climate and the lowest recorded annual rainfall may determines vegetation.

~400 MB of data reduced to ~1KB

Computing ET from First Principles

$$ET = \frac{\Delta Rn + \rho_a c_p (\delta q) g_a}{(\Delta + \gamma(1 + g_a/g_s))\lambda_v}$$

Penman–Monteith (1964)

ET = Water volume evapotranspired ($m^3\ s^{-1}\ m^{-2}$)

Δ = Rate of change of saturation specific humidity with air temp.($Pa\ K^{-1}$)

λ_v = Latent heat of vaporization (J/g)

R_n = Net radiation ($W\ m^{-2}$)

c_p = Specific heat capacity of air ($J\ kg^{-1}\ K^{-1}$)

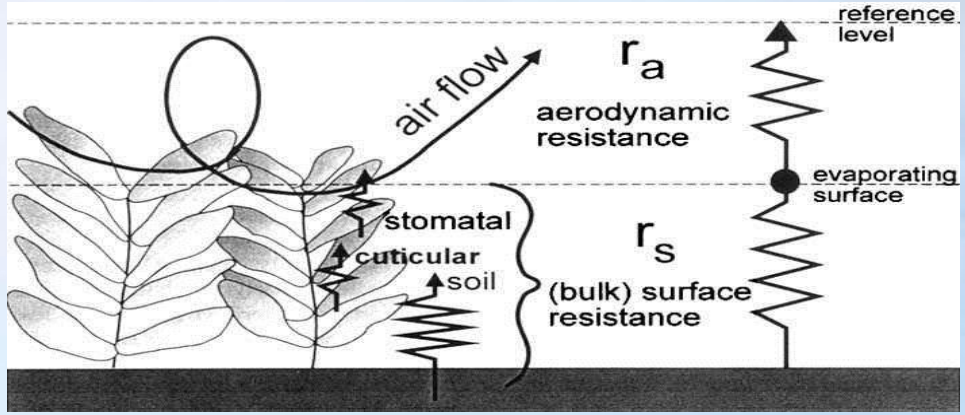
ρ_a = dry air density ($kg\ m^{-3}$)

δq = vapor pressure deficit (Pa)

g_a = Conductivity of air (inverse of r_a) ($m\ s^{-1}$)

g_s = Conductivity of plant stoma, air (inverse of r_s) ($m\ s^{-1}$)

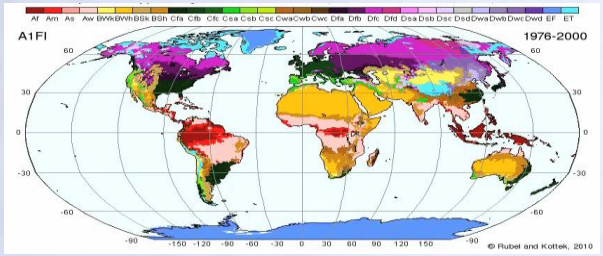
γ = Psychrometric constant ($\gamma \approx 66\ Pa\ K^{-1}$)



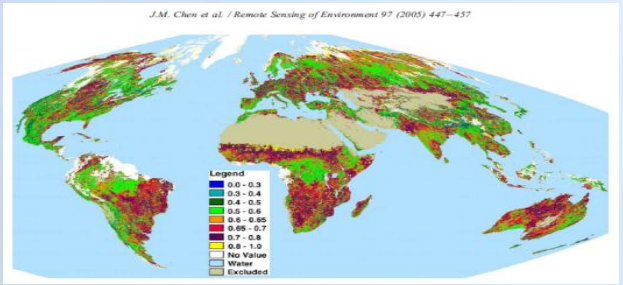
Estimating resistance/conductivity across a catchment can be tricky

- Lots of inputs : big reduction
- Some of the inputs are not so simple
- Many have categorical dependencies

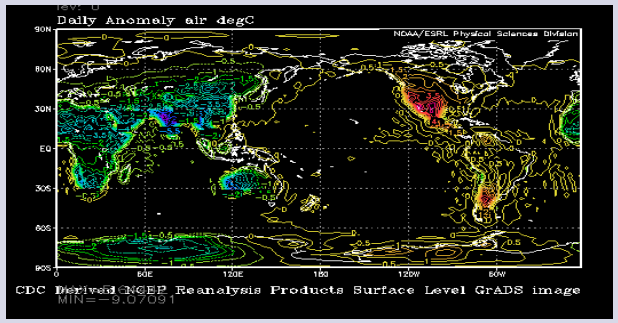
Computing ET from Imagery, Sensors and Field Data



Climate classification
~1MB (1file)

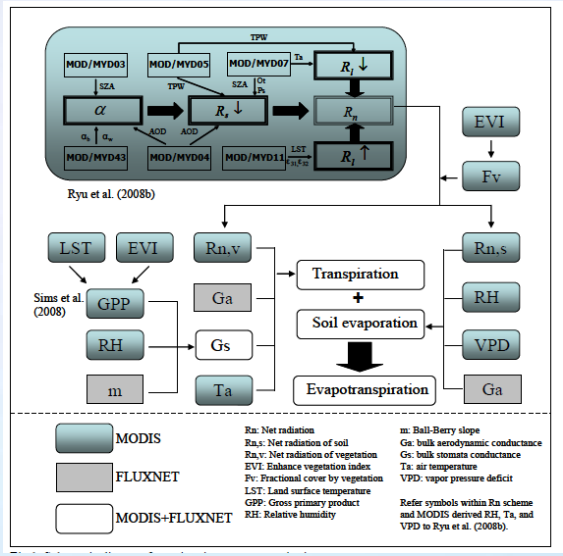


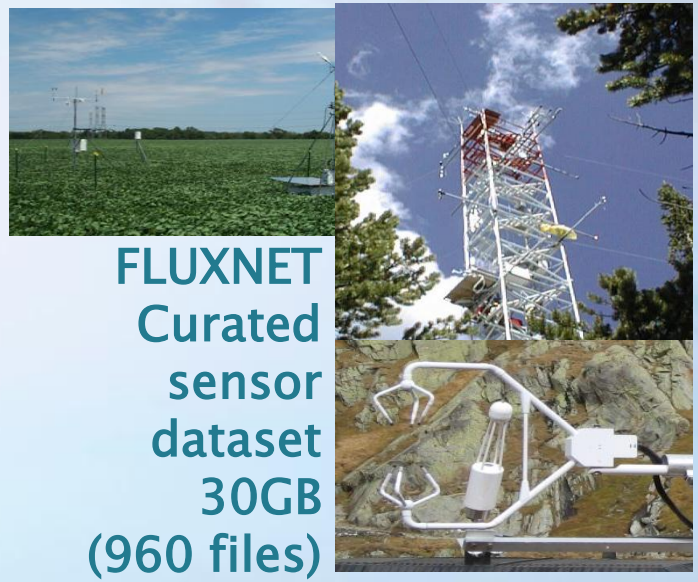
Vegetative clumping
~5MB (1file)



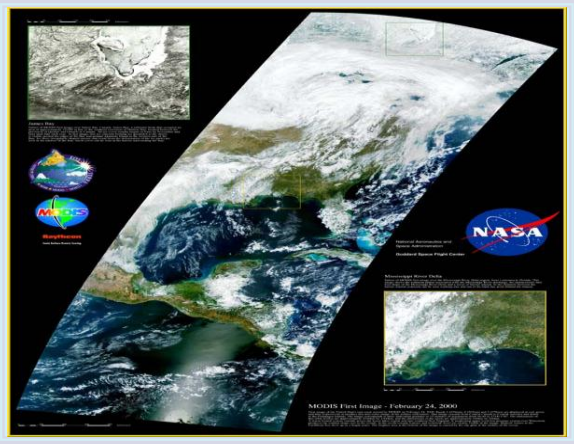
NCEP/NCAR ~100MB
(4K files)

Not just a simple matrix computation due to dry region leaf/air temperatures differences, snow cover, leaf area fill, temporal upscaling, gap fill, biome conductance lookup, C3/C4 plants, etc etc





FLUXNET
Curated
sensor
dataset
30GB
(960 files)



NASA MODIS imagery archives
5 TB (600K files) for 10 US years

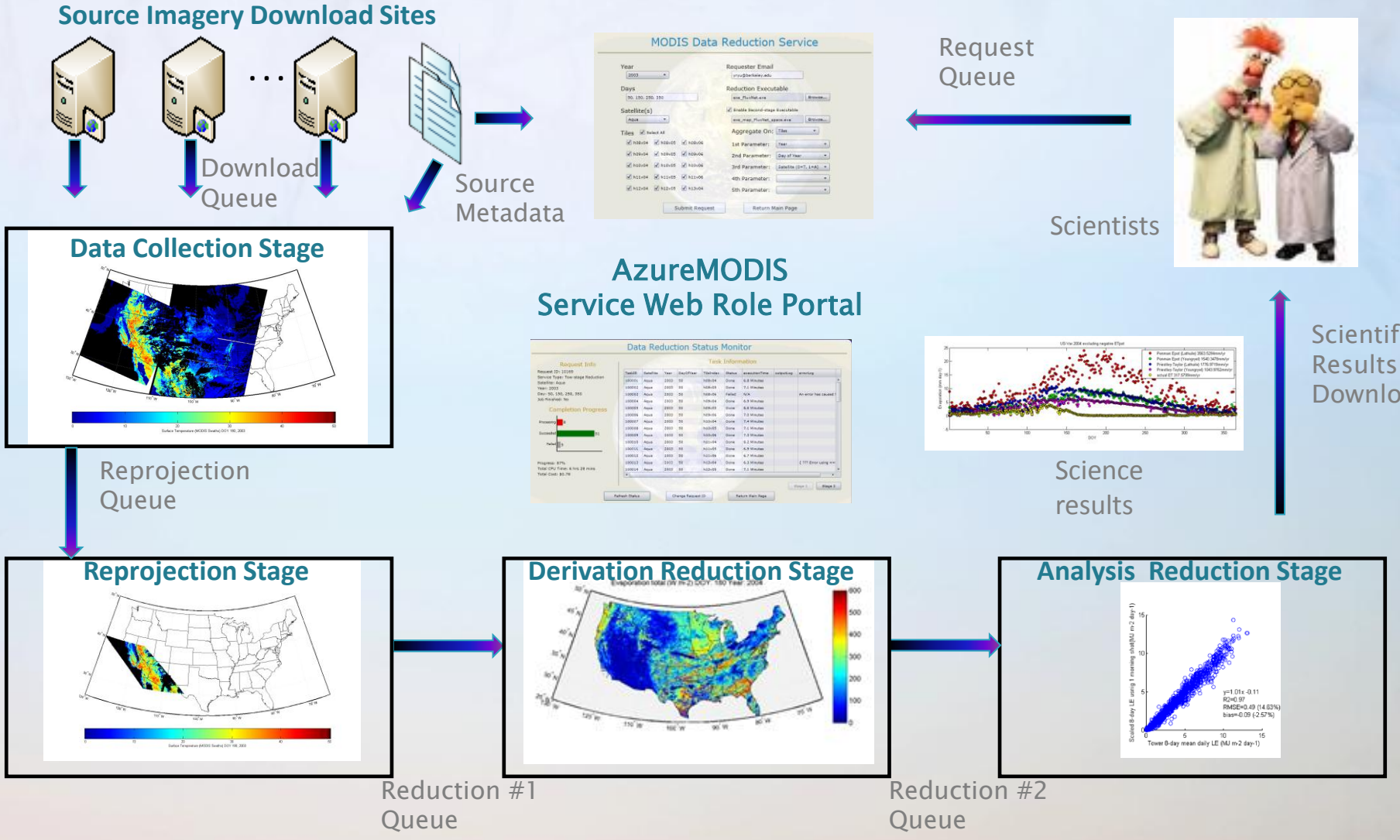
FLUXNET
curated field dataset
2 KB (1 file)

Satellites: MODISAzure

Behind every cloud is another cloud.
Judy Garland

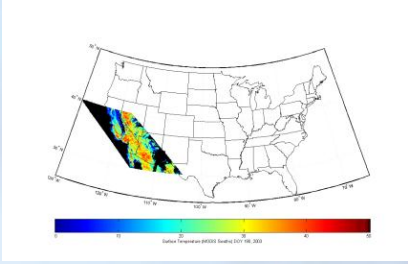
MODISAzure: Four Stage Image Processing Pipeline

- Data collection stage
  - Downloads requested input tiles from NASA ftp sites
  - Includes geospatial lookup for non-sinusoidal tiles that will contribute to a reprojected sinusoidal tile
- Reprojection stage
  - Converts source tile(s) to intermediate result sinusoidal tiles
  - Simple nearest neighbor or spline algorithms
- Derivation reduction stage
  - First stage visible to scientist
  - Computes ET in our initial use
- Analysis reduction stage
  - Optional second stage visible to scientist
  - Enables production of science analysis artifacts such as maps, tables, virtual sensors



http://research.microsoft.com/en-us/projects/azure/azuremodis.aspx

Determining What to Download



- Each product is either swath or sinusoidal projection
  - Sinusoidal are ready to use
  - Groups of swath products must be reprojected to create a sinusoidal tile
- NASA publishes a geometadata information for the two Terra and Aqua satellites
- For each 5 minute swath data file (or granule) on the ftp site there is a corresponding geometa file containing: DayNightFlag indicating day, night or both; corner point latitude/longitude; bounding coordinates
- We ingested all files (288 per day * 10 years * 2 satellites) into a SQL database then paged the information into our Azure ScanTimeList and GeoMeta Tables
- The dayScanTimeList in the ScanTimeList table identifies all swath source file precursors for a given sinusoidal tile and drives the download and reprojection

M*D04	Aerosol
M*D05	Precipitable water
M*D06	Cloud
M*D07	Temperature, ozone
MCD43B*	Albedo
M*D11	Surface temperature
M*D15	LAI
MOD13A2	Vegetation Index
MCD12Q1	Land Cover
MOD44B	Veg. Contig. Fields

#Attributes	PartitionK	RowKey	Timestamp	betweenScanTimeList	dayOfYear	dayScanTimeList	hIndex	nightScanTimeList	satellite	vIndex	year
Terra_2003_160		h00v07	2/10/2010 7:33		160	2220/2355/	0	1005/1010/1145/	Terra	0	2003

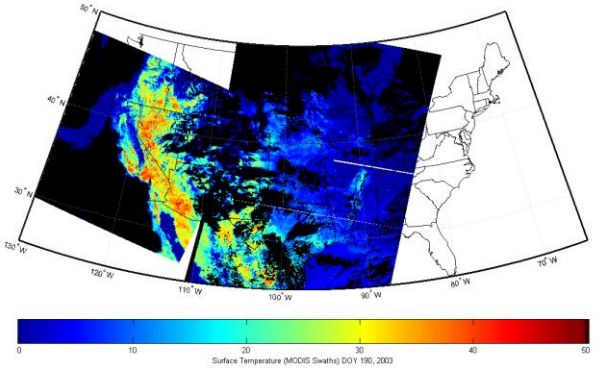
ftp://ladsweb.nascom.nasa.gov/geoMeta/README

Tiling: Do Scientists have to become Computer Scientists?

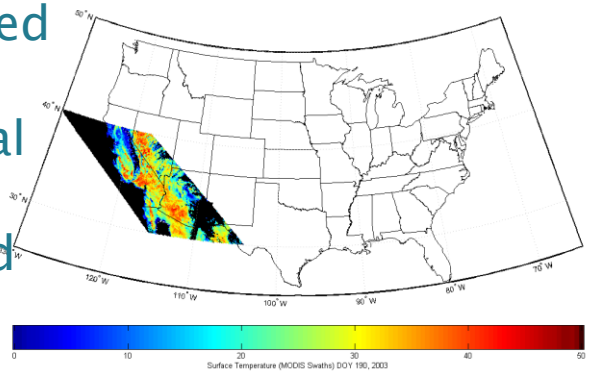
- Reprojection
  - Converts one geo-spatial representation to another.
  - Example: latitude-longitude swaths converted to sinusoidal cells.
- Spatial resampling
  - Converts one spatial resolution to another.
  - Example is converting from 1 KM to 5 KB pixels.
- Temporal resampling
  - Converts one temporal resolution to another.
  - Example is converting from daily observation to 8 day averages.
- Gap filling
  - Assigns values to pixels without data either due to inherent data issues such as clouds or missing pixels.
- Masking
  - Eliminates uninteresting or unneeded pixels.
  - Examples are eliminating pixels over the ocean when computing a land product or outside a spatial feature such as a watershed.

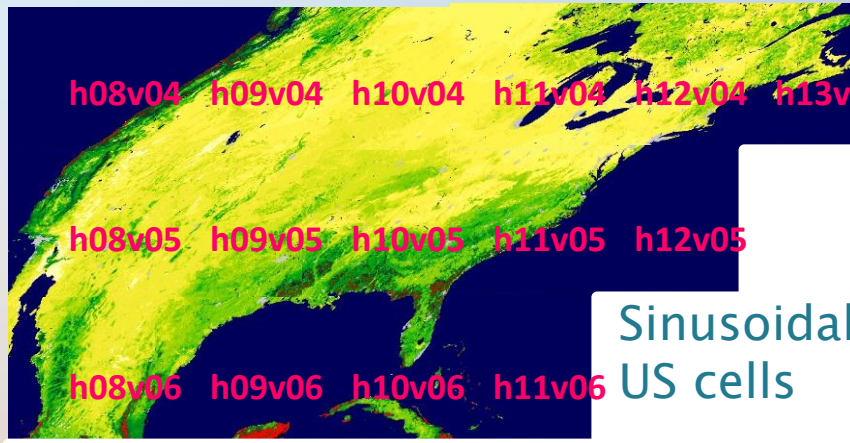
Grunge means you're doing science

Source Data (Swath format)



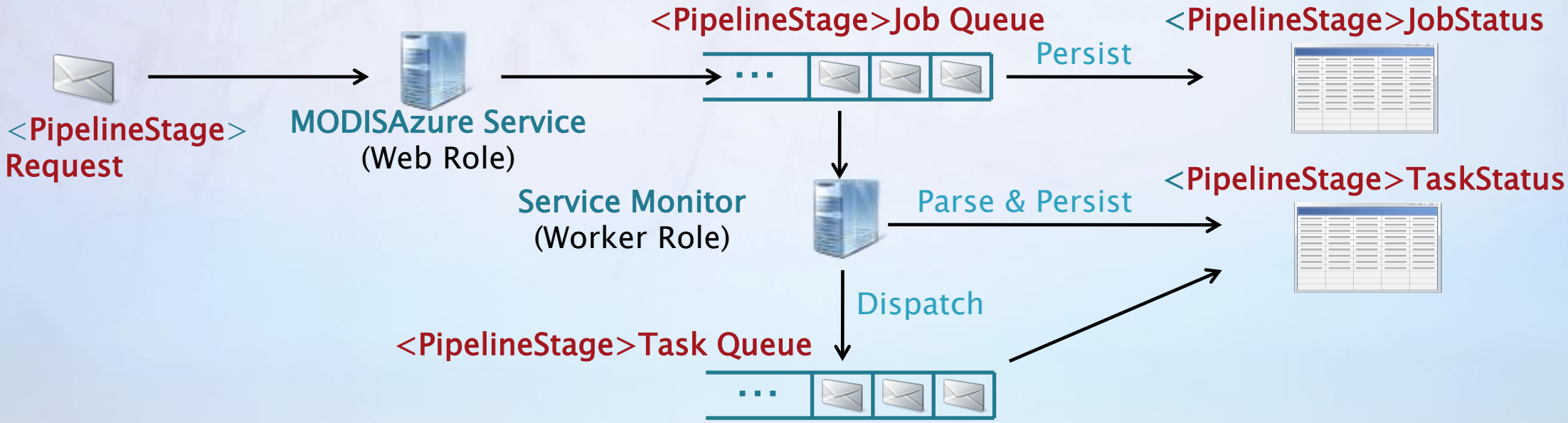
Reprojected Data (Sinusoidal format – equal land area pixel)





Sinusoidal US cells

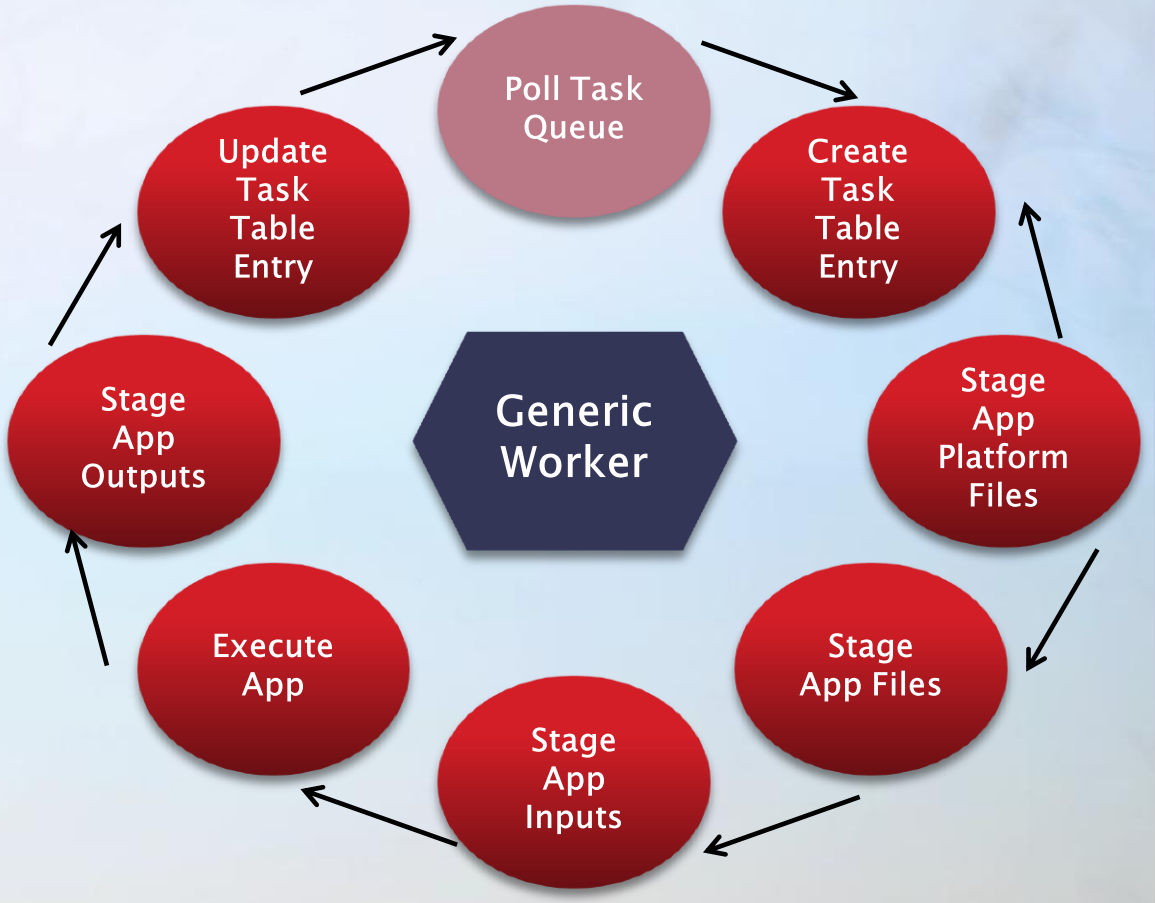
MODISAzure: Architectural Big Picture (1/2)



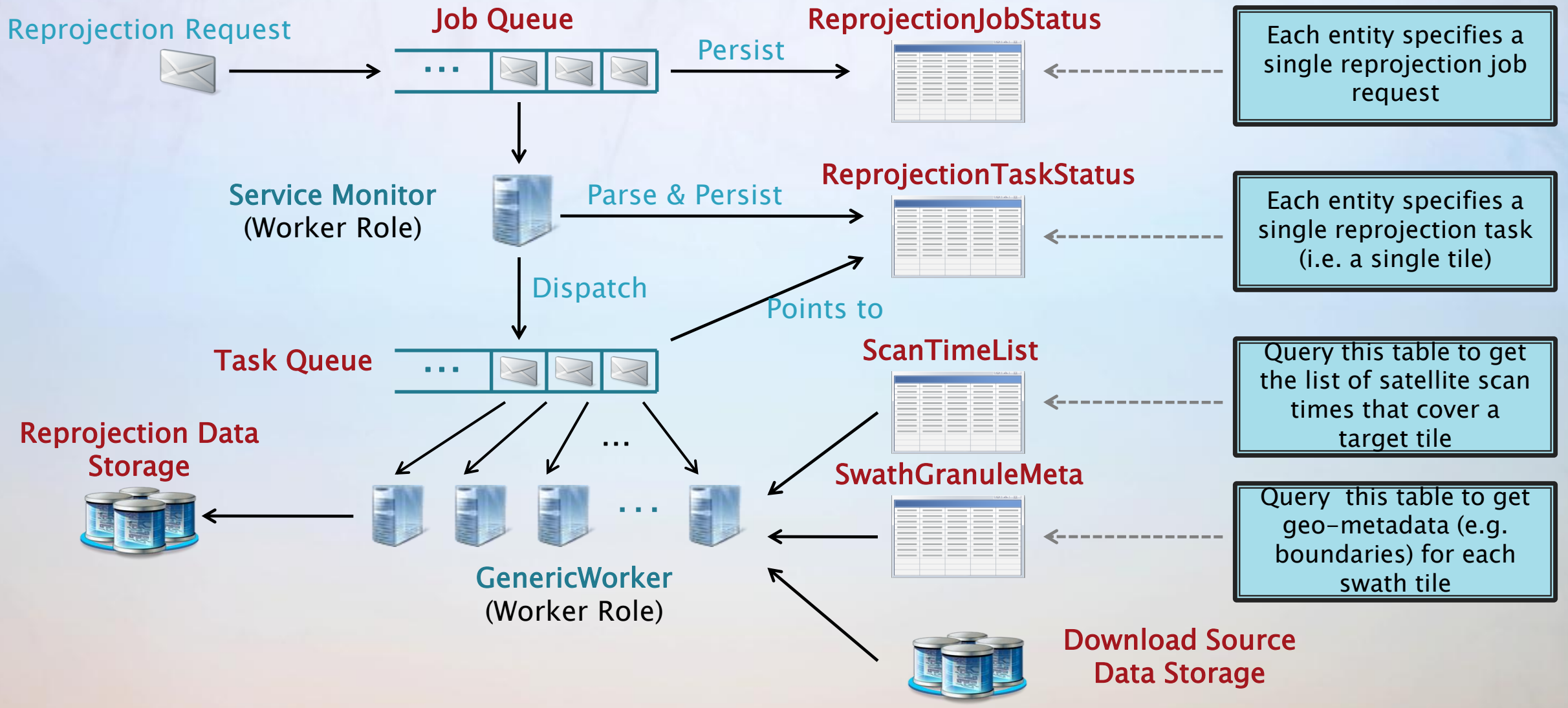
- **ModisAzure Service** is the Web Role front door
  - Receives all user requests
  - Queues request to appropriate Download, Reprojection, or Reduction Job Queue
- **Service Monitor** is a dedicated Worker Role
  - Parses all job requests into tasks – recoverable units of work
  - Execution status of all jobs and tasks persisted in Tables

All Work is Done by a GenericWorker

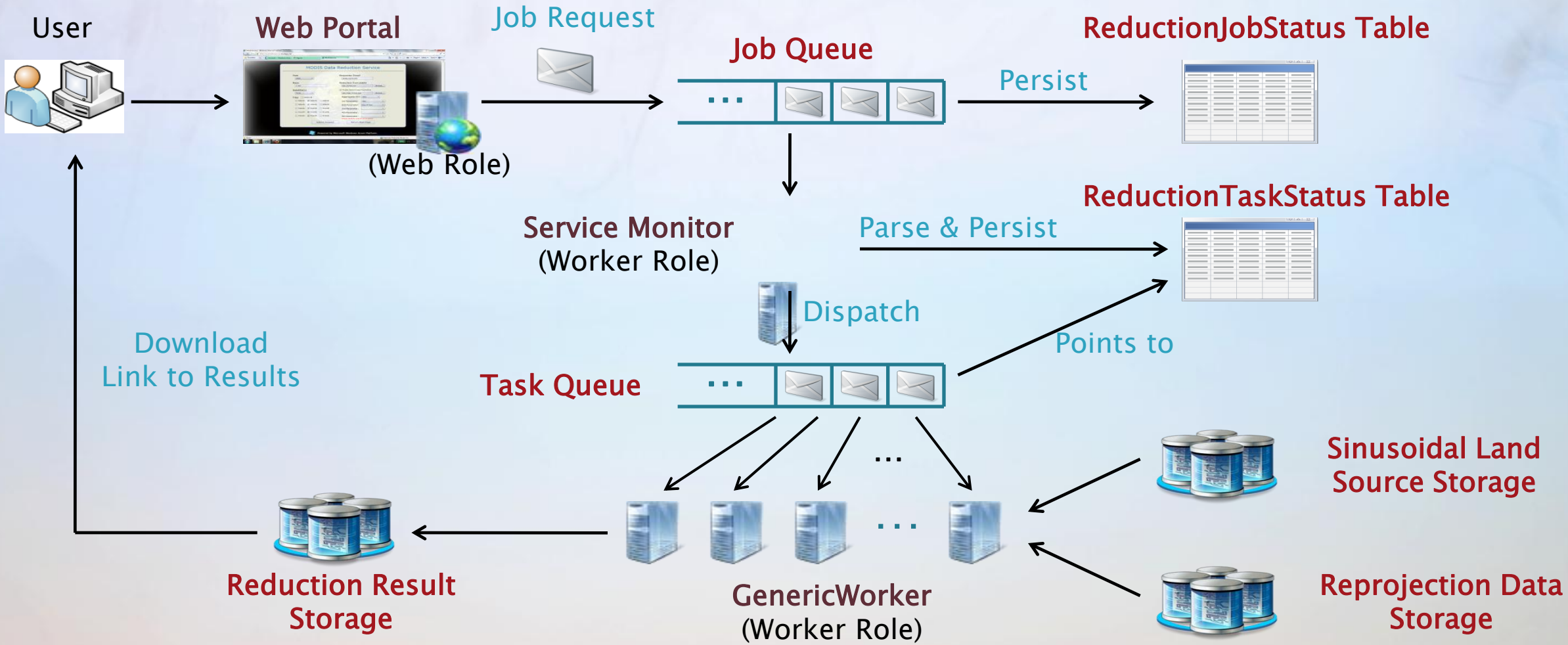
- Manages application sandbox
  - Ensures all application binaries such as the MatLab runtime are installed for “known” application types
  - Stages all input blobs from Azure storage to local files
  - Passes any marshalled inputs to uploaded application binary
  - Stages all output blobs to Azure storage from local files
  - Preserves any marshalled outputs to the appropriate Task table
- Manages all task status
  - Dequeues tasks created by the Service Monitor
  - Retries failed tasks 3 times
  - Maintains all task status
- Simplifies desktop development and cloud deployment



Reprojection Service

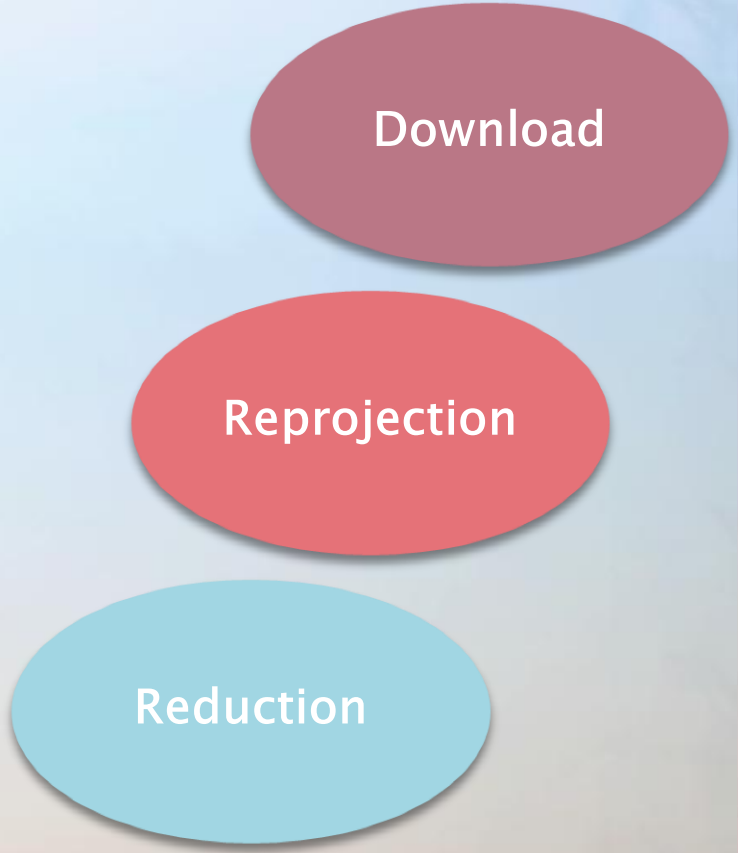


Reduction Service (Only One Stage Shown)



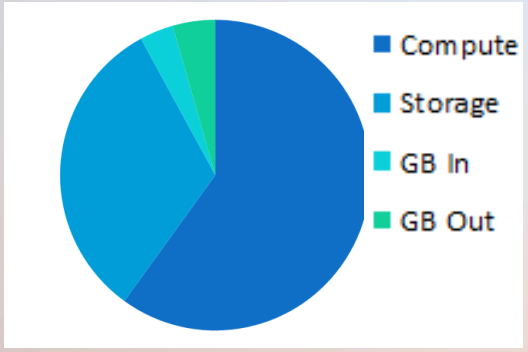
Pipeline Stage Interactions

- The Web Portal Role, Service Monitor Role and 5 Generic Worker Roles are deployed at most times
  - 5 Generic Workers are sufficient for reduction algorithm testing and development ($20/day)
  - Early results returned to scientist while deploying up to 93 additional Generic Workers; such a deployment typically takes 45 minutes
  - Deployment taken down when long periods of idle time are known
  - Heuristic for scaling number of Generic Workers up and down
- Download stage runs in the deep background in all deployed generic worker roles
  - IO, not CPU bound so no competition
- Reduction tasks that have available inputs run preferentially to Reprojection tasks
  - Expedites interactive science result generation
  - If no available inputs and a backlog of reprojection tasks, number of Generic Workers scale up naturally until backlog addressed and reduction can continue
  - Second stage reduction runs only after all first stage reductions have completed
- Reduction results can be downloaded following emailed link to zip file

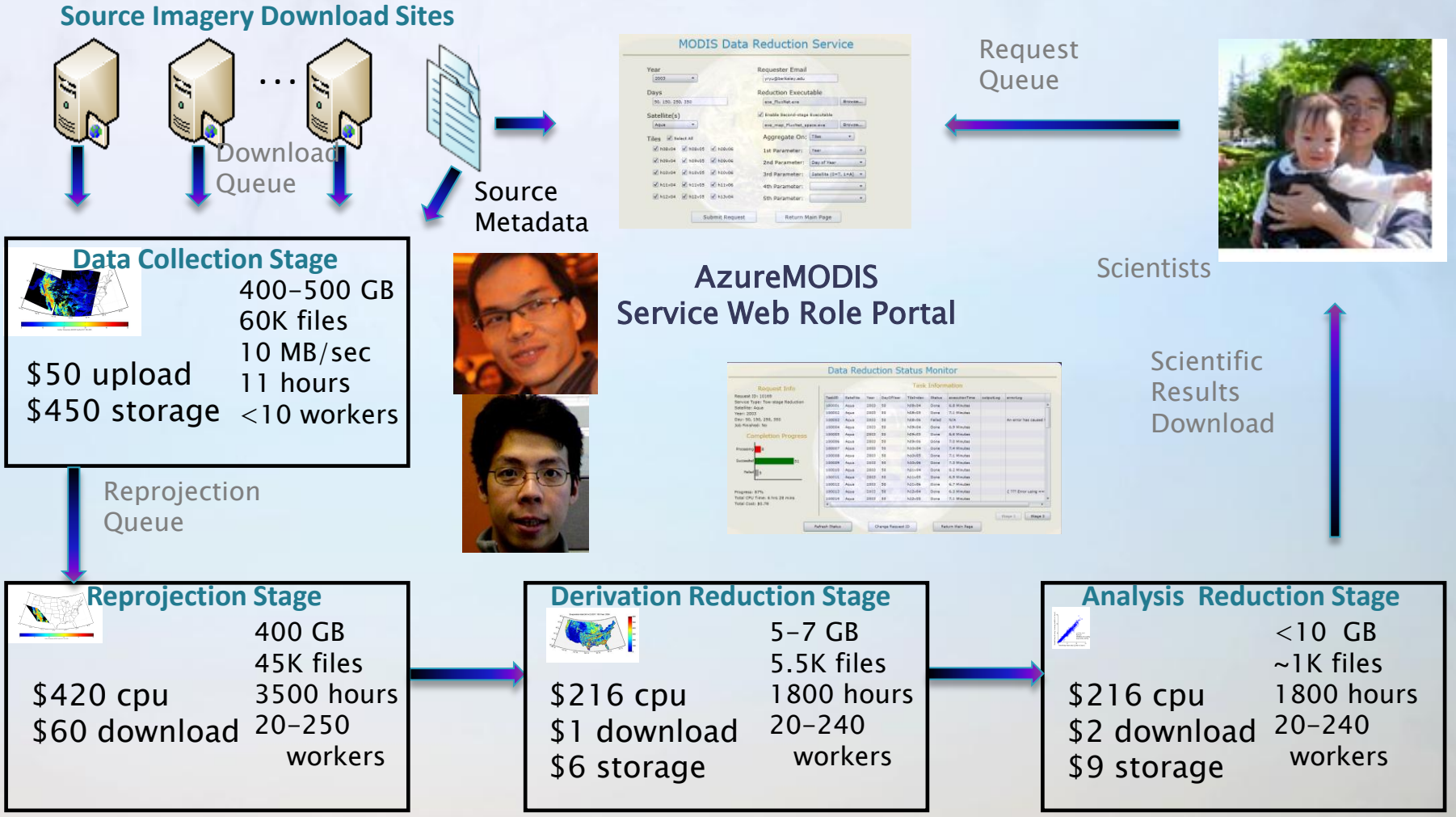


Costs for 1 US Year ET Computation

- Computational costs driven by data scale and need to run reduction multiple times
- Storage costs driven by data scale and 6 month project duration
- Small with respect to the people costs even at graduate student rates !



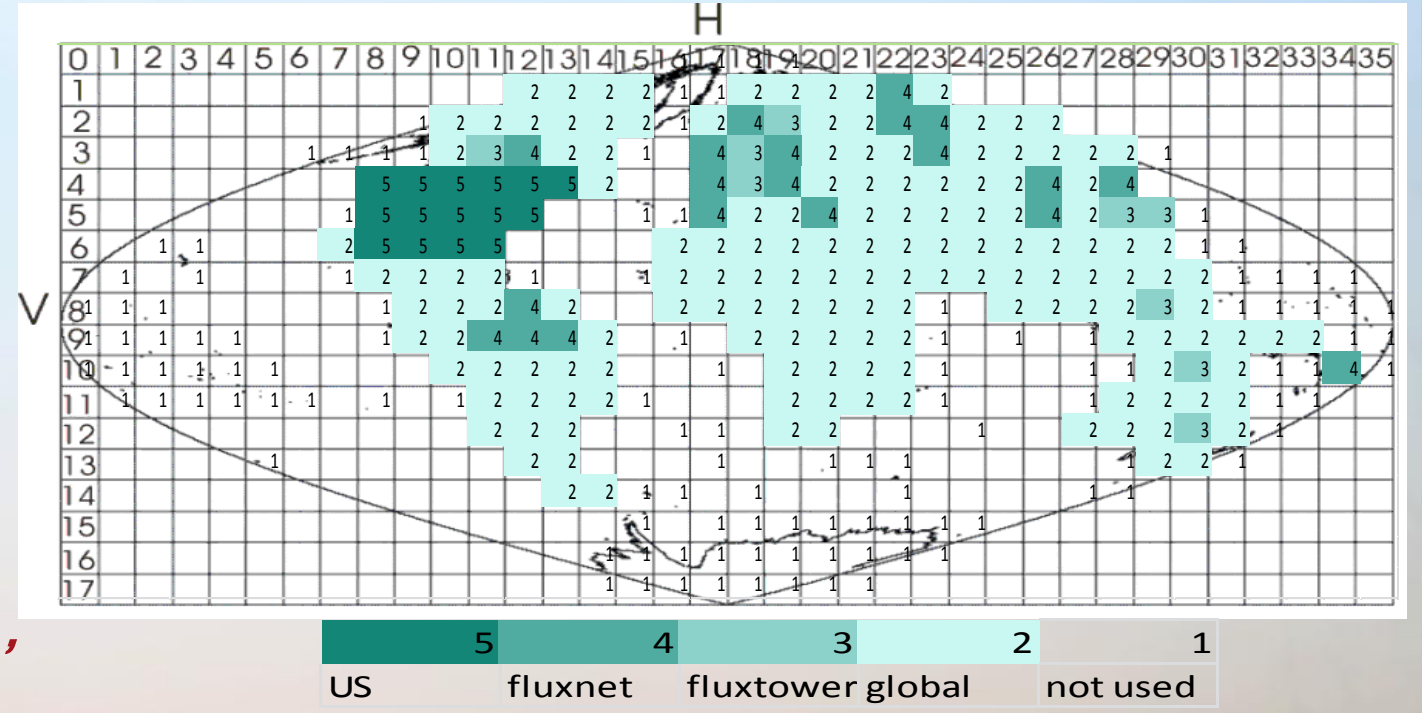
Total: $1420



Sizing the 3 year MODISAzure Global Computation

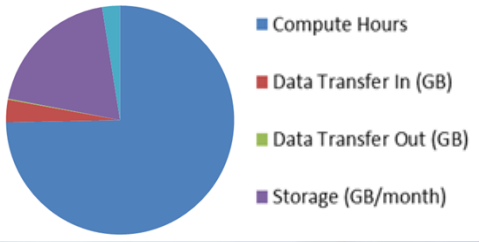
- 194 sinusoidal cells, each covers 1.2x1.2 KM or 11M 5 KM pixels)
- 1.06 M reprojected tiles and 40.5K source sinusoidal tiles
- 8 TB (>10 M files) downloaded from NASA ftp
- Not all files are downloaded or reprojected at first (3 rapid retries) attempt or actually available due to satellite outage, polar winter, missing tiles, etc etc.
- 55 NASA download days
- 150K reprojection compute hours
- 940 TB moved across Azure fabric
- 10 download result days (est) via IN2 bridge

15 seconds on the Cray Jaguar (1.75 PFLOPs) , but only if we could get the PB in

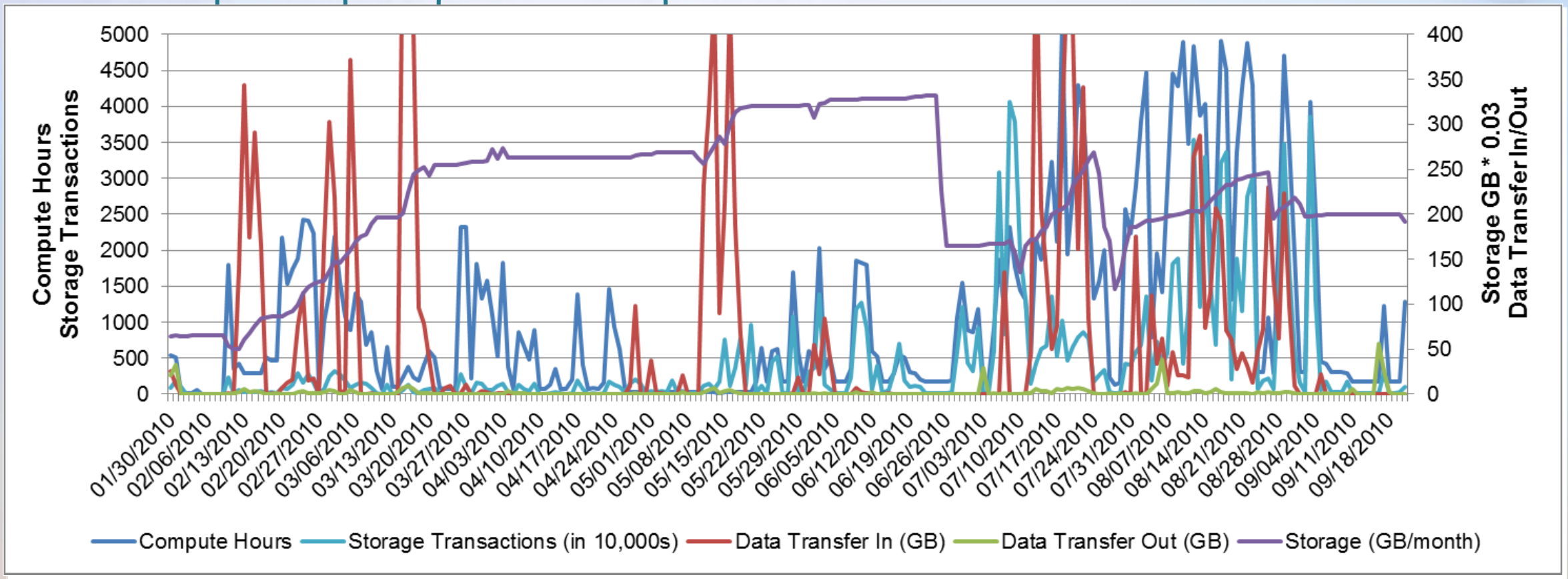


Agility

Cumulative MODISAzure billing ($39K)



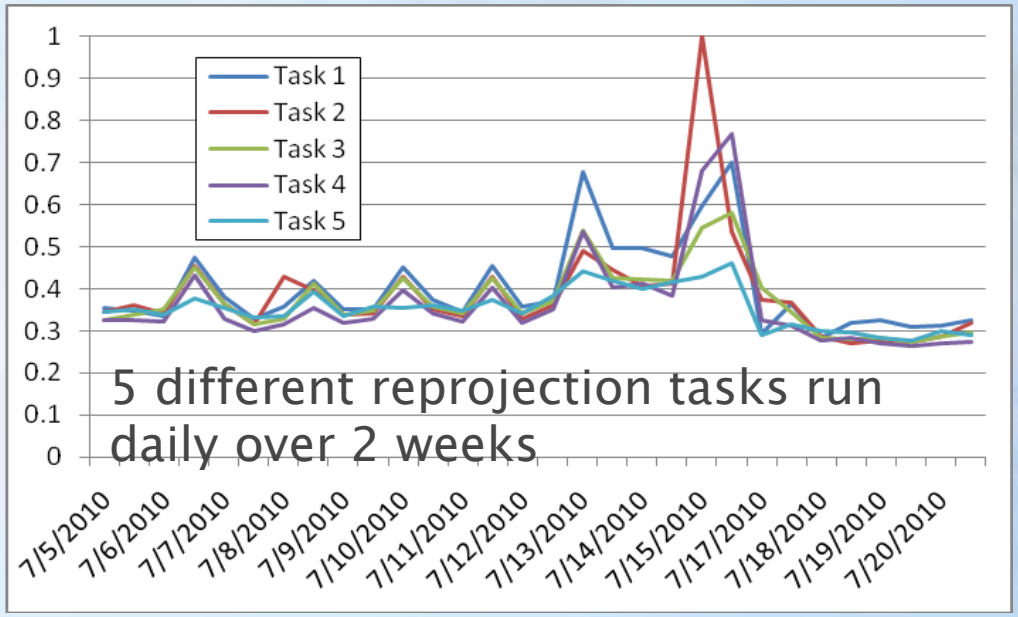
US years 3-10 | 3 FLUXNET years | Global scale lower resolution

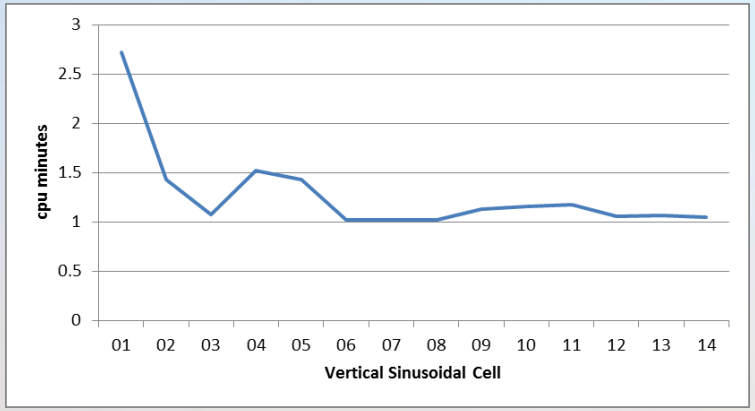


The computation changed over time while Azure just scaled

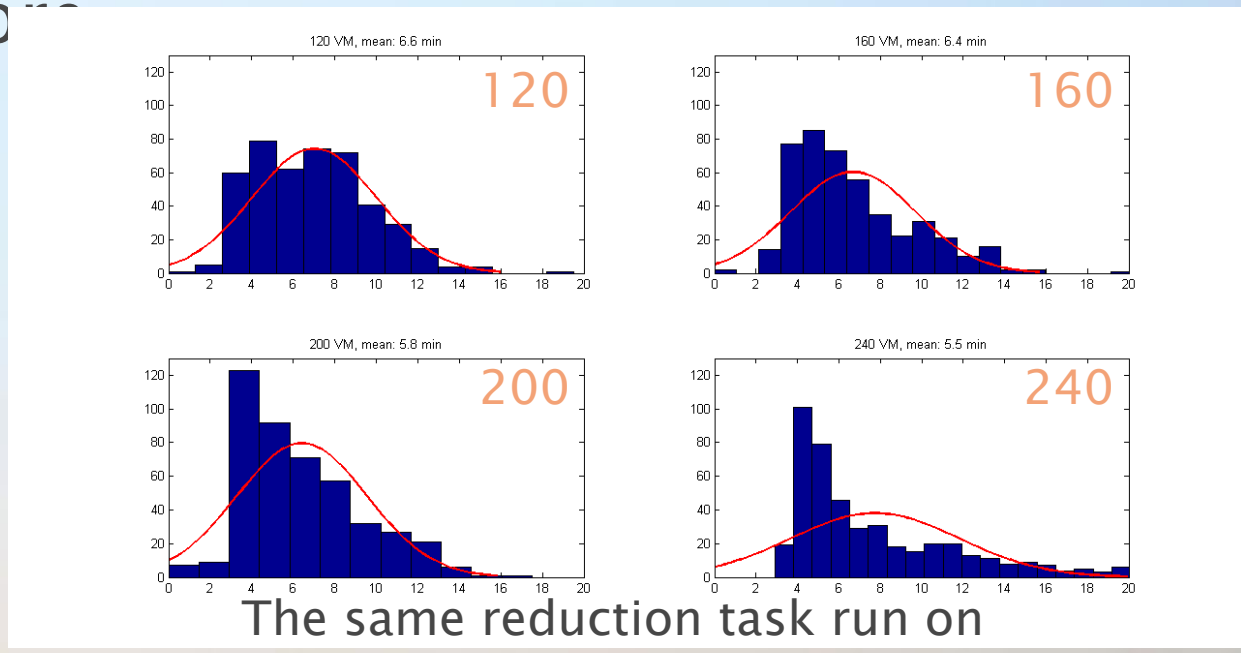
Predictability

- Performance varies over time: rerunning the same task gives different timings on different days
- Performance varies over space: satellites are over the poles more often





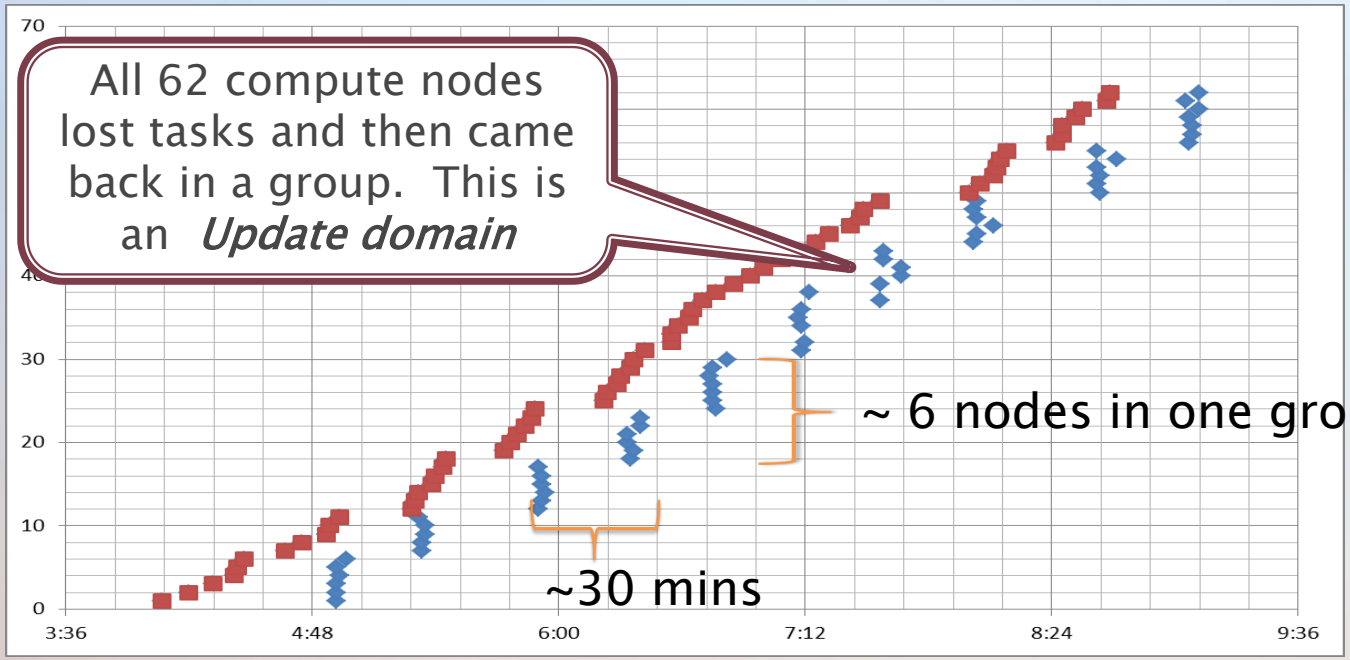
Average reprojection time (after algorithm improvements!) as a function of longitude



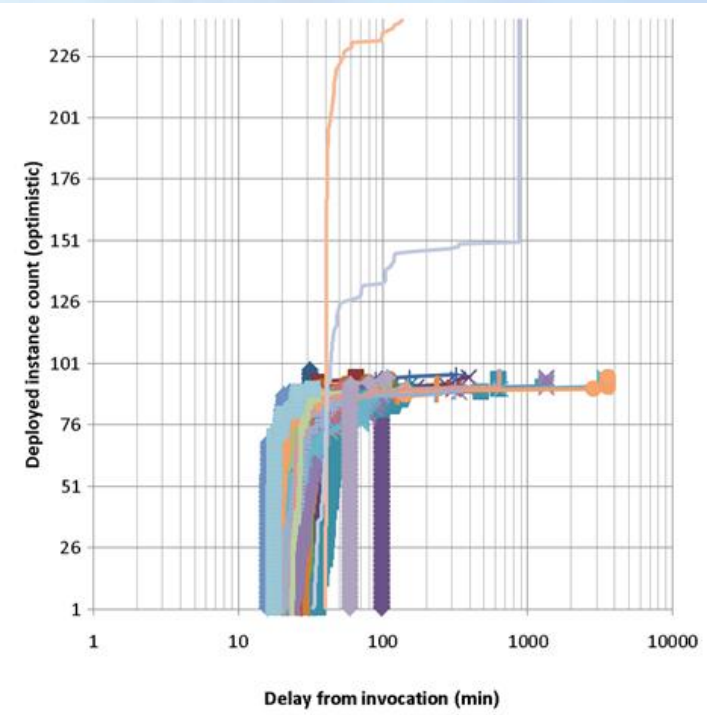
The same reduction task run on different numbers of VMs

Reliability

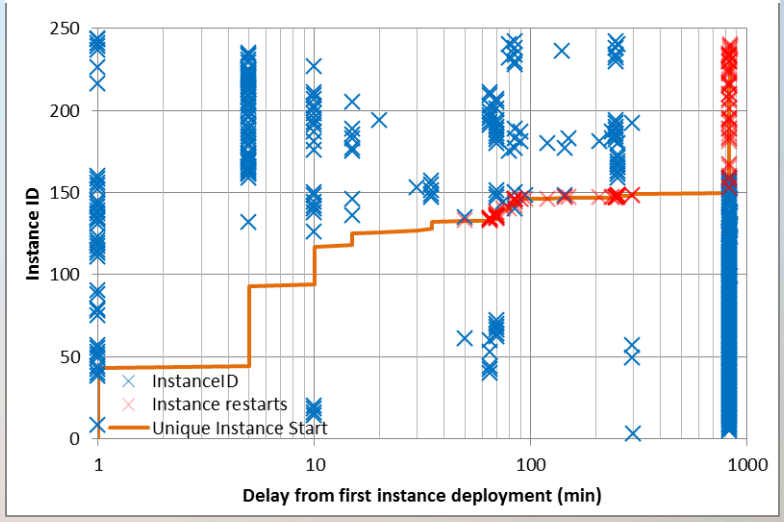
- Even with 99.999% reliability, bad things happen
  - 1-2 % of MODISAzure tasks fail but succeed on retry



Observed VM starts for 76-100 VMs



Worst case 250 VM start



From AzureBlast
http://research.microsoft.com/en-us/people/barga/faculty_summit_2010.pdf

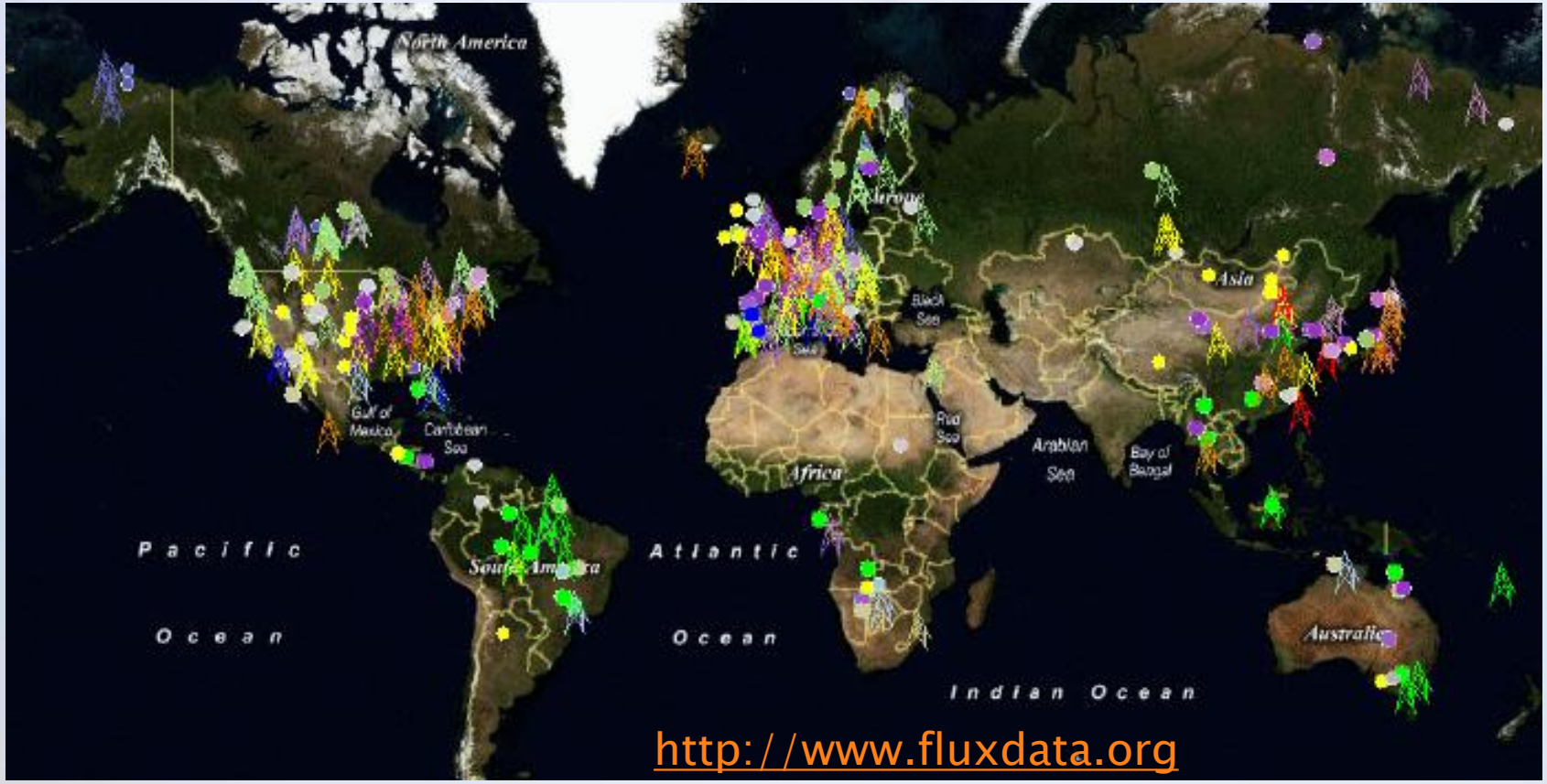
Maintainability

- Some “Early Adopter” artifacts
  - Generic worker sandbox
  - “dir” for blobs : need to have a parsable list, not just browse and many tools simply could not scale beyond O(50K) blobs
  - “downloader” for blobs : early SDK utility retired by end of CTP.
  - Slow upload (FEDEX disk is still “in plan”; IN2 connections helped download tremendously
- Can we move catalog and other tracking to SQL Azure for better scaling?
  - Current tracking database is 140 GB
  - Partitions naturally, but would mean $300/mo (external) charges.

Sensors: fluxdata.org

In wilderness I sense the miracle of life,
and behind it our scientific accomplishments fade to trivia.
Charles Lindbergh

FLUXNET: A Network of Networks

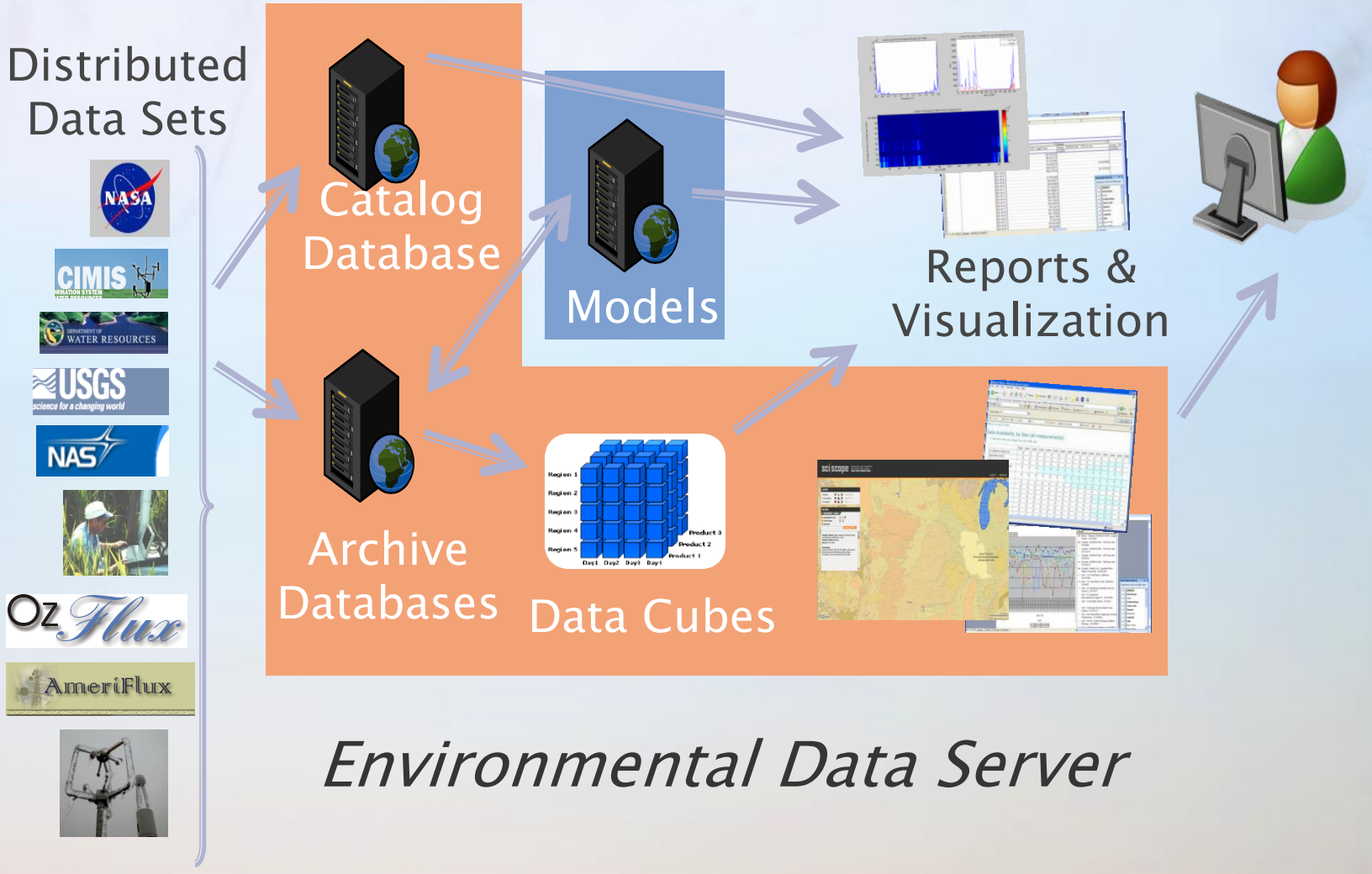


- 467 towers world wide
- 967 site-years of sensor data from 253 towers (~800K data points)
- ~20 sensor measurements per tower; 20 derived science variables
- 145 ancillary variables

- Original data set assembled and processed in 2007
- 20x larger than previous synthesis dataset
- Currently 85 paper teams with over 400 scientists

Initial Challenge: Connecting People and Data

- Leverages commercial technologies
  - SQL Server 2008 database
  - Sharepoint 2008 collaborative portal
- Cloud service with transparent desktop connection to common desktop tools such as Excel and MatLab

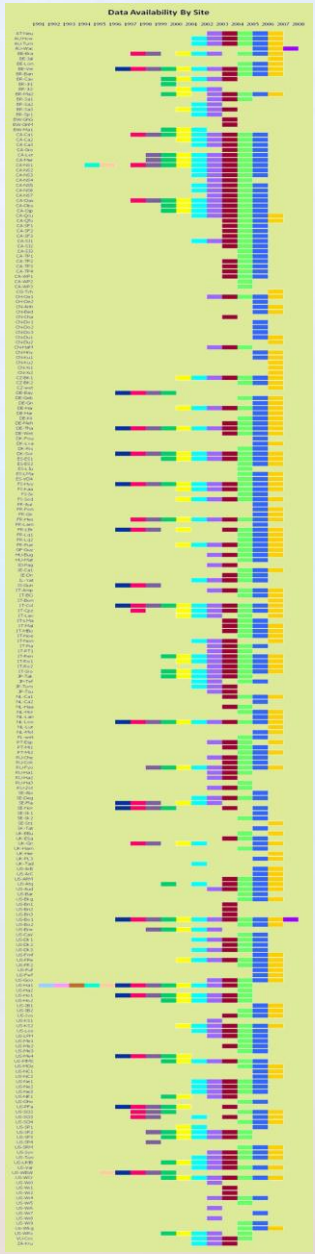


Data Set Assembly

- L0 (raw sensor output) take at 1-10 Hz stored at tower site
- L1 (sensor calibration applied) and L2 (initial science variable derivation) performed by the tower team; L2 is aggregated to 30 minute samples.
- L3 (quality assessment); L4 (gap fill) and L5 (additional science variable derivation) performed by common processing across the networks and/or by the regional network
- L4/L5 data flows from regional network to Europe for common processing and then to fluxdata.org data access portal
- Ancillary data (disturbances, tree rings, leaf chemistry, phenology) submitted to regional network or directly to fluxdata.org in common format and processed by fluxdata.org team
  - Analyses often require combining time series data with fixed, or nearly fixed ancillary data; ancillary data used as fixed property, time series, or event time window.

Database Schema Abstractions

- Sensor data, ancillary data and metadata
  - Normalized table structure simplifies adding variables and cube building
- Versioning and folder-like collections
  - Accommodate algorithm changes
  - Track derivations throughout processing Define and track analysis “working set”
- Namespace translations
  - Data assembly traverses different repositories each with own name space
  - Some repositories encode metadata in variable name space
  - Units conversions when aggregating temporally or spatially

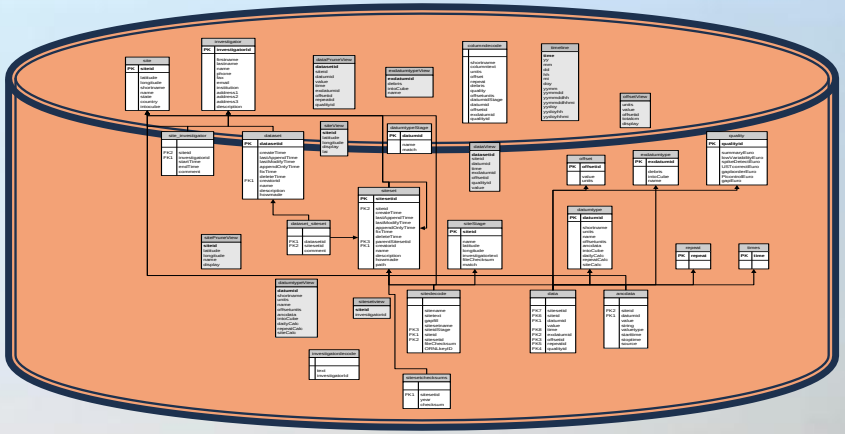


Analysis Tools
Excel, MatLab, R, ArcGIS

Map Mashups

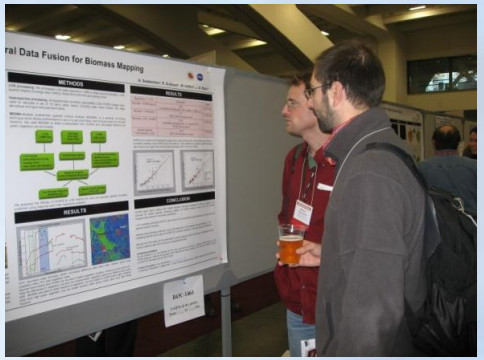
Spreadsheet data summaries

CSV, MatLab and Excel file import/export



All access layers share the same abstractions

Current Challenge: Connecting Data, People and Publications



- Data browsing, curation, and archive is only the beginning
- Increase data sharing, data reuse and, most important, science learning
- Publish data with papers
  - Capture contributor information, fair-use criteria, and acknowledgments
  - Capture analysis artifacts and connect them to both data source and paper learnings
  - Capture then current data corrections, gap-fills, quality assessment inferences for subsequent forensics

Global Convergence in the Temperature Sensitivity of Respiration at Ecosystem Level

Miguel D. Mahecha[1,2*], Markus Reichstein[1], Nuno Carvalhais[1,3], Gitta Lasslop[1], Holger Lange[4], Sonia I. Seneviratne[2], Rodrigo Vargas[5], Christof Ammann[6], M. Altaf Arain[7], Alessandro Cescatti[8], Ivan A. Janssens[9], Mirco Migliavacca[10], Leonardo Montagnani[11,12], Andrew D. Richardson[13]

[1]Max Planck Institute for Biogeochemistry, 07745 Jena, Germany.

[2]Institute for Atmospheric and Climate Science, ETH Zurich Universitätsstrasse 16, 8092 Zurich, Switzerland.

[3] Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

[4]Norsk Institutt for skog og landskap, N-1431 Ås, Norway.

[5] Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA.

[6]Agroscope ART, Federal Research Station, Reckenholzstr. 191, CH-8046 Zurich, Switzerland

[7]McMaster Centre for Climate Change, McMaster University, Hamilton, Ontario, Canada.

[8]European Commission, Joint Research Center, Institute for Environment and Sustainability, Ispra, Italy.

[9]Department of Biology, University of Antwerpen, Universiteitsplein 1, 2610 Wilrijk, Belgium.

[10]Remote Sensing of Environmental Dynamics Laboratory, DISAT, University of Milano-Bicocca, Milano, Italy.

[11]Servizi Forestali, Agenzia per l'Ambiente, Provincia Autonoma di Bolzano, Bolzano, Italy.

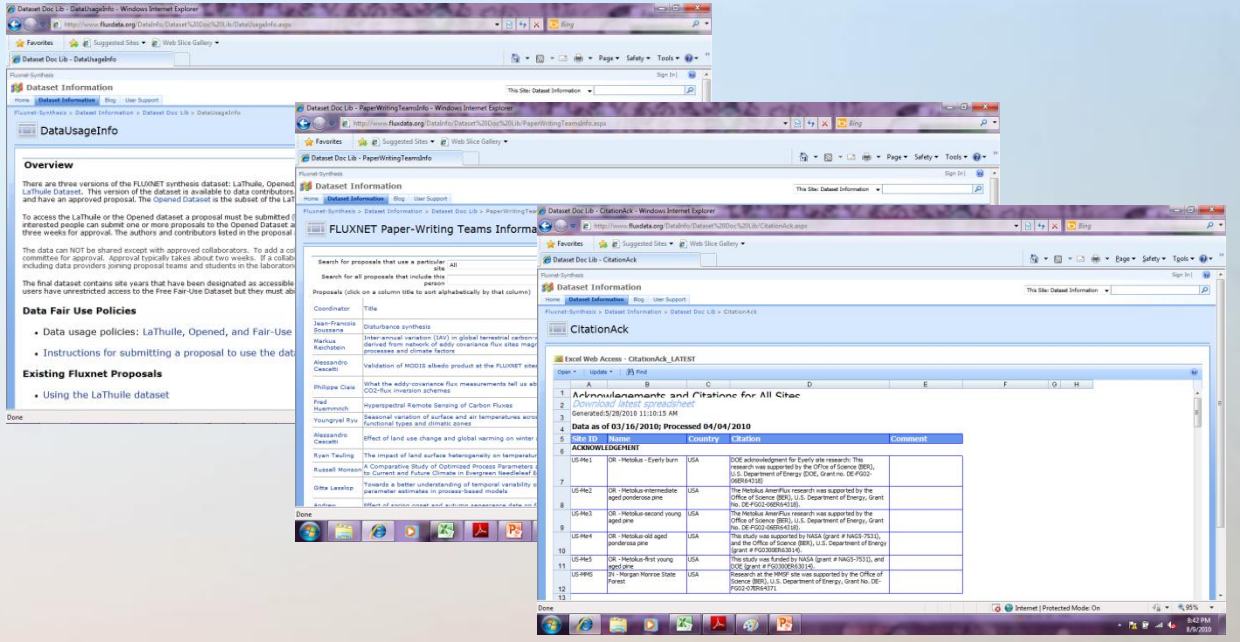
[12]Faculty of Sciences and Technologies, Free University of Bozen-Bolzano, Piazza Università 1, 39100, Bolzano, Italy.

[13]Department of Organismic and Evolutionary Biology, Harvard University, HUH, 22 Divinity Avenue, Cambridge, MA 02138, USA.

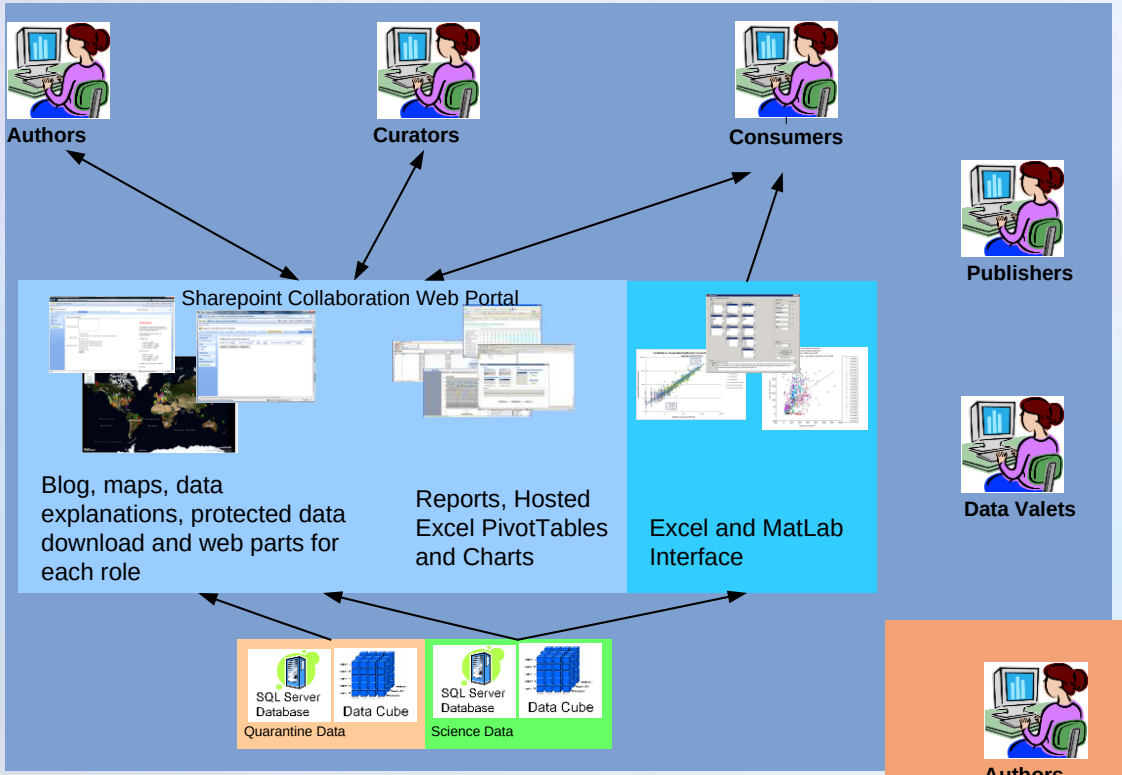
*To whom correspondence should be addressed; E-mail: mmahecha@bgc-jena.mpg.de.

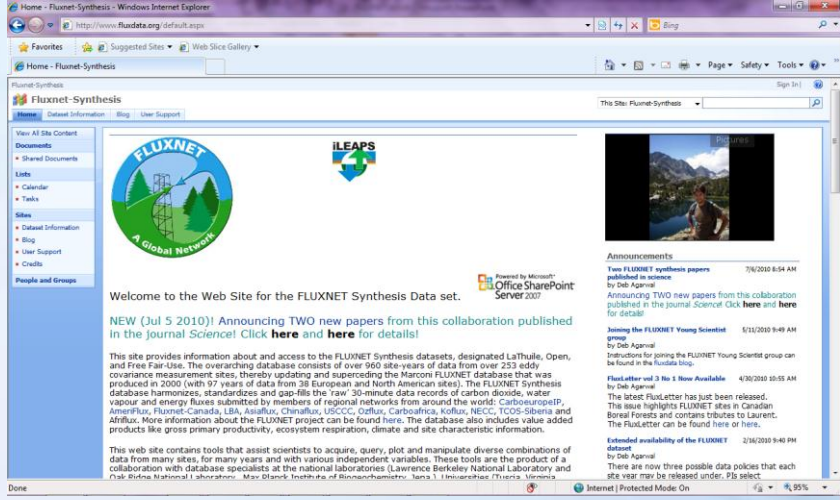
Data Centric Publication Aware Collaboration Portal

- Enables a “FLUXNET collaboration” author
  - Ability to track the then current tower data contributors, data processors and science committee for each paper
  - Tower owners can discover and track papers that use their data
  - Enumerated citations and acknowledgements
- 3 level data sharing policy
  - Opt in for tower owners
- More pictures !



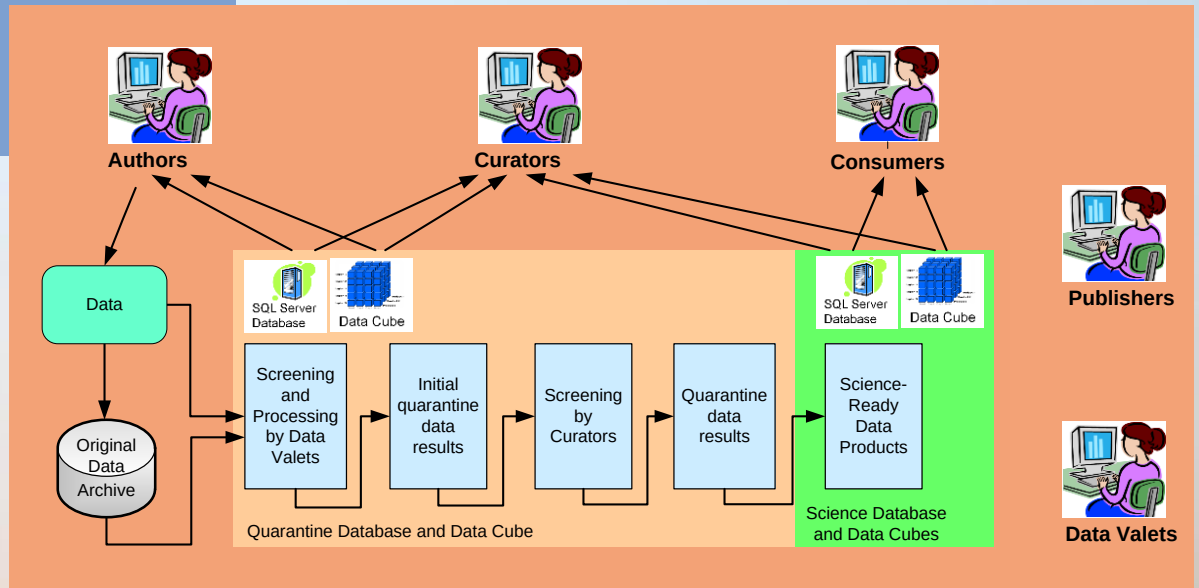
Enabling Virtual Conversations





Collaboration Centric View

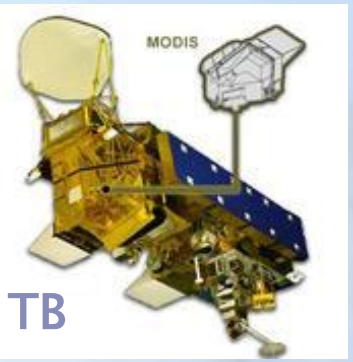
Data Centric View



Summary and Prepublication Results

Baby, it's the beginning of a great adventure.
Lou Reed

Fluxdata Learnings

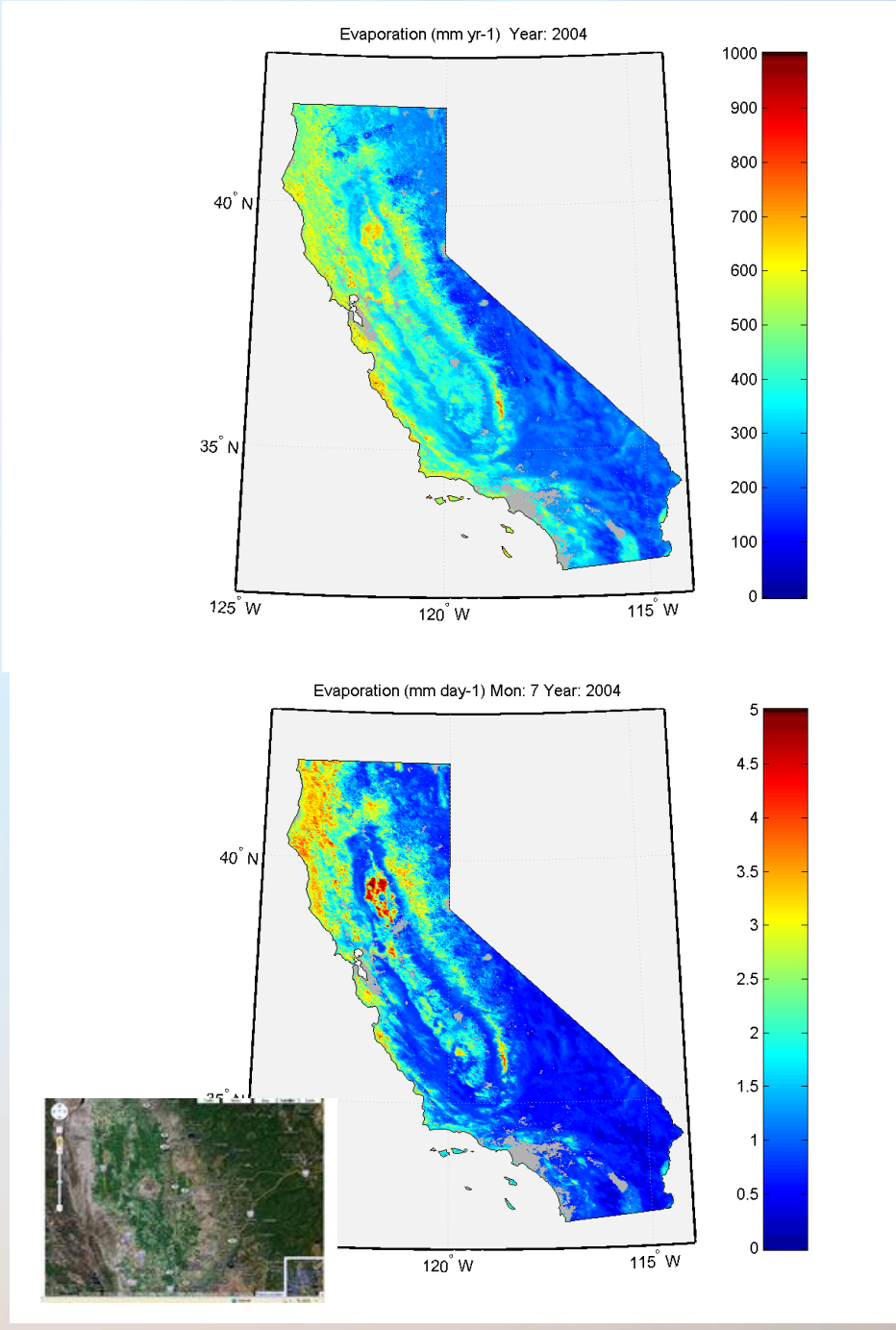


- Provenance and trust widely varies
  - Data acquisition, early processing, and reporting ranges from a large government agency to individual scientists.
  - Smaller data often passed around in email; big data downloads can take days (if at all)
  - Opaque safe-deposit boxes and storage lockers prevail today
- Data sharing concerns and patterns vary
  - Open access followed by (non-repeatable and tedious) pre-processing
  - True science ready data set but concerns about misuse, misunderstanding particularly for hard won data.
- Computational tools differ.
  - Not everyone can get an account at a supercomputer center
  - Very large computations require engineering (error handling)
  - Space and time aren't always simple dimensions

Science happens when PBs, TBs, GBs, and KBs can be mashed up simply

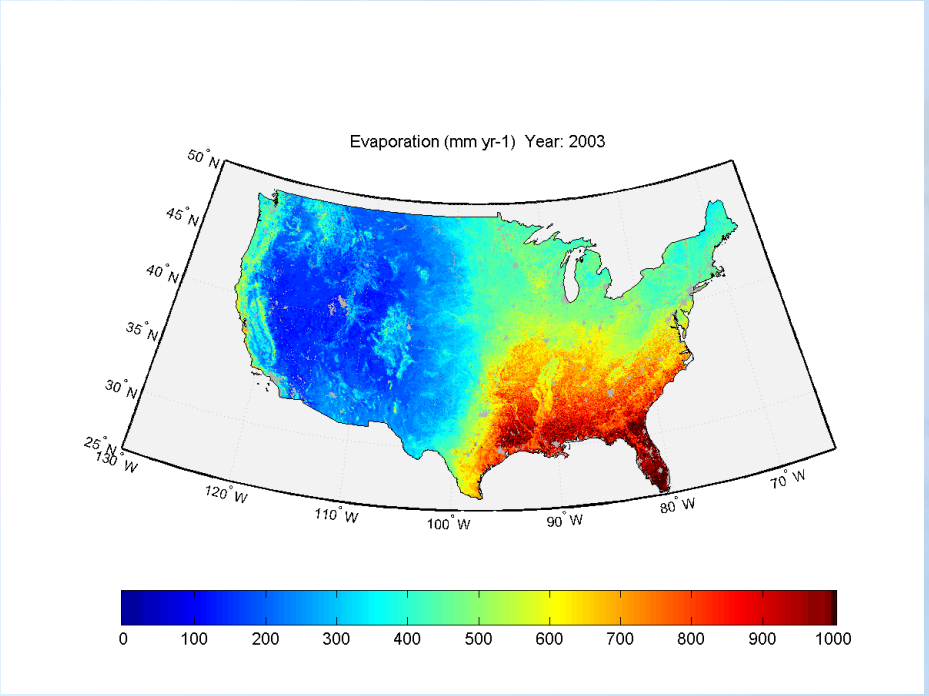
Azure Learnings

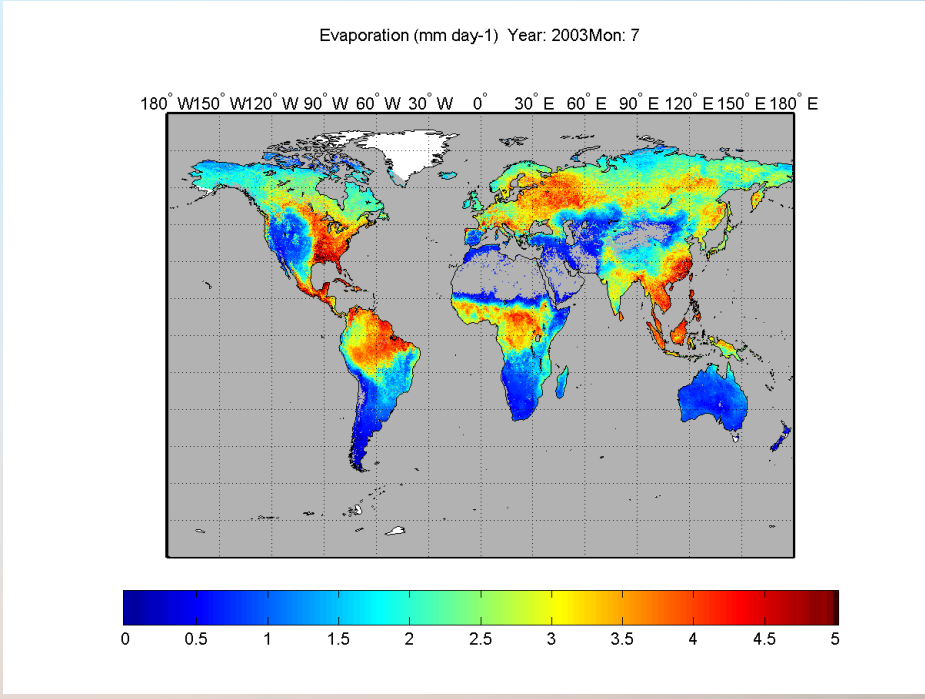
- Putting all your eggs in the cloud basket means watching that basket
  - Cloud scale resources often mean you still manage small numbers of resources: 100 instances over 24 hours = $288 even if idle
  - Where is the long term archive for any results ?
- Azure is a rapidly moving target and unlike the Grid
  - Commercial cloud backed by large commercial development team
  - Current target applications are mid-range or smaller – MODISAzure is currently at the fringe
- Scaling up requires additional work as understanding even a 0.01% failure rate is time consuming
  - Bake in the faults for scaling and resilience
  - Bake in the catalog for end:end reconciliation of sources and results



eScience Learnings

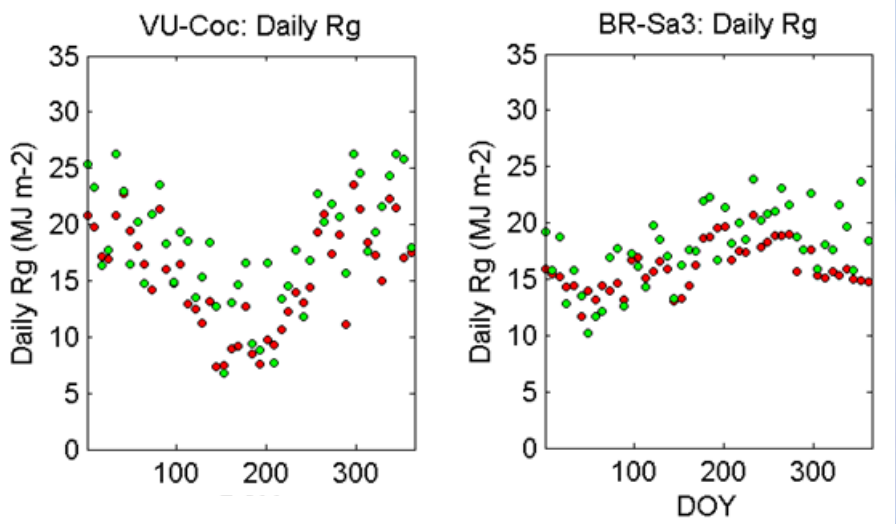
- Lowering the barriers to use remote sensing data can enable science
  - NASA makes the data accessible, not science ready
  - At AGU 2009, we learned that a cloud service that just made on-demand jpg mosaics would help tremendously
- Science and algorithm debugging benefit from the same infrastructure as both need to scale up and down
  - Debugging an algorithm on the desktop isn't enough – you have to debug in the cloud too
  - Whenever running at scale in the cloud, you must reduce down to the desktop to understand the results
- Scaling up means expanding the science
  - California, New England, and Florida are different
  - Boreal forests, savannahs, fertilization practices differ across the globe

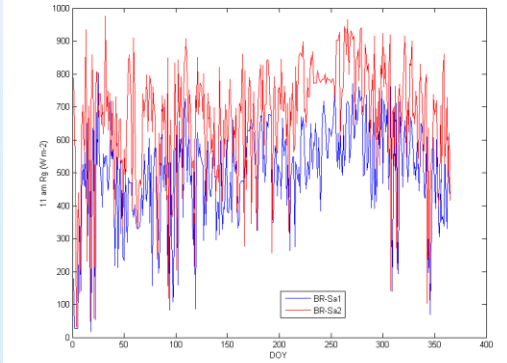


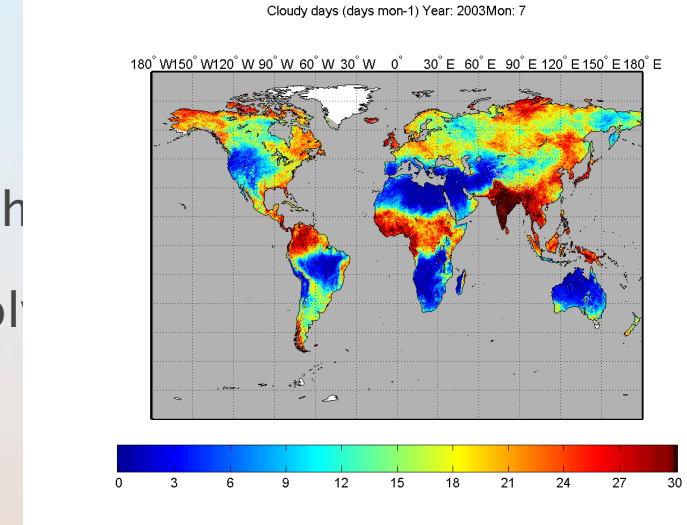


4th Paradigm Learnings

- Developing concrete plans for validation, sensitivity analysis, and mining a large computation prior to having results in hand is tricky
  - Precedents break down when scaling 100x or more
  - Sub-discipline familiarity a good start – our initial plans centered on FLUXNET tower data
  - Large sanity check aggregates a good start – our carbon fixation is in range of the literature estimates
  - Watershed aggregate comparison in the US crossed disciplines, length, and time scales as well as introducing yet different grunge.
- "Everybody knows" local knowledge plays a big role
  - Citizen science opportunity is anecdotal rather th quantified voting
  - Machine learning seems possible, but likely invol categorical geospatial subdivisions and some science cross checking







Current computation underestimates solar radiation in the tropics.
Some of the comparison sites had to be discarded due to systematic drift of sensors used.
And that data set has been used by others and passed the FLUXNET QA/QX.

4th Paradigm Challenges Forward

- Since the dominant cost is people, how can we generalize the compute infrastructure to a wider class of computations?
  - Would Dryad/HPC/LINQ be faster, easier, more maintainable ?
- How should we proceed to understanding our global computation and related other computations well enough to improve such a computation over the next few years ?
  - Are there aggregate approaches such as computing statistics rather than values then statistics to reduce the overall computation requirements?
  - What should we do about the science factors we omitted such as elevation changes ?
  - What is the role for machine learning ? How can we engage?

Acknowledgements

Microsoft Research
- Dan Reed
- Tony Hey
- Dennis Gannon
- David Heckerman
- Nelson Araujo
- Dan Fay
- Jared Jackson
- Wei Liu
- Jaliya Ekanayake
- Simon Mercer
- Yogesh Simmhan
- Michael Zyskowski

Berkeley Water Center, University of California, Berkeley, Lawrence Berkeley Laboratory
- Deb Agarwal
- Dennis Baldocchi
- Jim Hunt
- Monte Goode
- Susan Hubbard
- Keith Jackson
- Rebecca Leonardson
- Carolyn Remick

University of Virginia
- Marty Humphrey
- Norm Beekwilder
- Jie Li

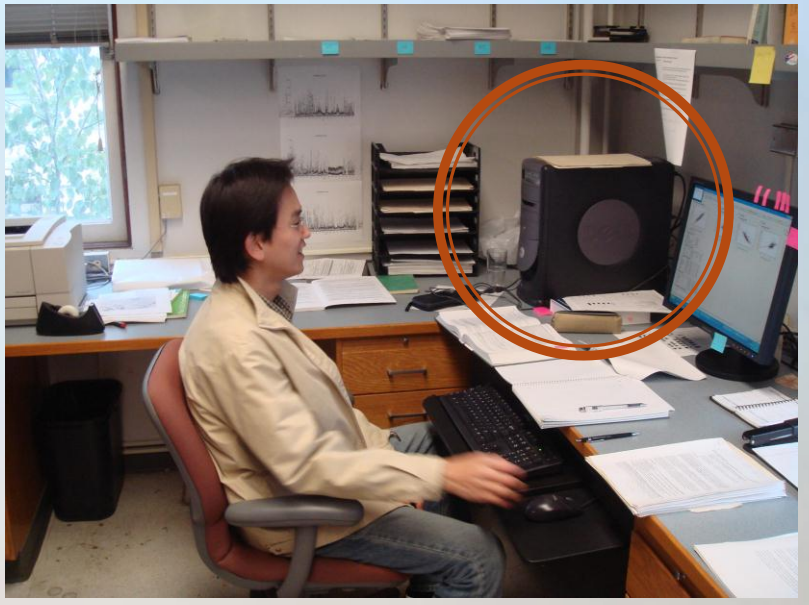
Indiana University
- You-Wei Cheah

Fluxnet Collaboration
- Dennis Baldocchi
- Youngryel Ryu
- Dario Papale (CarboEurope)
- Markus Reichstein (CarboEurope)
- Alan Barr (Fluxnet-Canada)
- Bob Cook
- Dorothea Frank
- Susan Holladay
- Hank Margolis (Fluxnet-Canada)
- Rodrigo Vargas

Ameriflux Collaboration
- Beverly Law
- Tom Boden
- Mattias Falk
- Tara Hudiburg (student)
- Hongyan Luo (postdoc)
- Gretchen Miller (student)
- Lucie Ploude (student)
- Andrew Richardson
- Andrea Scheutz (student)
- Christophe Thomas

http://www.fluxdata.org

Youngryel was lonely with 1 PC



From Youngryel Ryu, 2010